Conditional Reversible Work method: a novel approach to obtain pair potentials for coarse-grained simulation of soft matter

TECHNISCHE UNIVERSITÄT DARMSTADT

Vom Fachbereich Chemie
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades eines
Doktor rerum naturalium (Dr. rer. nat.)

genehmigte
Dissertation

vorgelegt von

Emiliano Brini,
Dottore Magistrale (Dott. Mag.)

aus Bologna, Italien

Referent: Prof. Dr. Nico F. A. van der Vegt
Korreferent: Prof. Dr. Florian Müller-Plathe
Tag der Einreichung: 13. September 2012
Tag der mündlichen Prüfung: 5. November 2012

Darmstadt, 2012
D17

Contents

Summary v

Zusammenfassung ix

Acknowledgements xiii

1 Introduction 1
1.1 Bibliography 4

Publications 5

2 Systematic coarse graining methods for soft matter simualtion: a review 7
2.1 Introduction 8
2.2 CG interaction potentials 9
2.3 CG bonded potentials 10
2.4 CG nonbonded potentials 12
2.4.1 Parametrized methods 12
2.4.2 Derived Methods 15
2.5 Applications 17
2.5.1 Apolar molecular liquids 17
2.5.2 Polar molecular liquids - water 19
2.5.3 Polymers - Polystyrene 22
2.5.4 Ionic solutions 24
2.6 Concluding remarks 27
2.7 Acknowledgment 27
2.8 Bibliography 28

3 CRW methods for molecular CG applications 33
3.1 Introduction 34
3.2 Conditional Reversible Work - CRW 36
3.2.1 Implicit solvent CRW potentials 38

3.3 Computational details . . . 39
3.4 Results and discussion . . . 40
3.4.1 Mapping schemes . . . 40
3.4.2 Hexane . . . 41
3.4.3 Toluene . . . 43
3.4.4 Comparison between CRW and IBI models of toluene . . . 45
3.5 Discussion and Conclusions . . . 48
3.6 Acknowledgements . . . 50
3.7 Bibliography . . . 51
3.8 Supplementary informations of chapter 3 . . . 53
3.8.1 Symmetry of the CG interaction potential with respect to the original symmetry of the molecules . . . 53

4 Chemically transferable coarse-grained potentials from conditional reversible work calculations 55
4.1 Introduction . . . 56
4.2 Coarse Graining using the Conditional Reversible Work Method . . . 57
4.2.1 Thermodynamic Cycle: Potential of Mean Force Method . . . 58
4.2.2 Direct Calculation by Free Energy Perturbation . . . 58
4.3 Models . . . 60
4.4 Computational Details . . . 60
4.4.1 MD simulations with the OPLS model . . . 61
4.4.2 CG potentials . . . 62
4.4.3 MD simulations with the CG models . . . 62
4.5 Results and Discussion . . . 63
4.5.1 Non bonded interaction potentials . . . 63
4.5.2 Representability and thermodynamic transferability . . . 64
4.5.3 Chemical transferability . . . 67
4.6 Discussion and Conclusions . . . 70
4.7 Acknowledgement . . . 73
4.8 Bibliography . . . 74

5 Thermodynamic transferability of coarse-grained potentials for polymer-additive systems 77
5.1 Introduction . . . 78

5.2 Methodology 80
5.2.1 Conditional Reversible Work method 80
5.2.2 Excess chemical potentials 82
5.2.3 Thermodynamic analysis 82
5.3 Models 84
5.3.1 United-atom model 84
5.3.2 Coarse-grained model 85
5.4 Computational details 86
5.4.1 CRW 86
5.4.2 Excess chemical potentials 87
5.4.3 MD simulation 87
5.5 Results and discussion 87
5.5.1 CG potentials 88
5.5.2 ECPs of additives in polystyrene melts 89
5.5.3 Structure 91
5.5.4 Entropy and enthalpy contributions to ECPs 93
5.6 Conclusions 95
5.7 Bibliography 96
5.8 Supplementary informations of chapter 5 99
5.8.1 Half atom groups and free energy calculations 99

6 Limits in the applicability of Conditional Reversible Work pair potential 103
6.1 Introduction 104
6.2 Conditional Reversible Work Method 105
6.2.1 Sampling of the interaction in bulk liquid 105
6.3 Models 106
6.3.1 Atomistic models 106
6.3.2 Coarse-grained model 106
6.4 Computational details 107
6.4.1 CRW calculations 107
6.4.2 Comparison between atomistic and CG models 108
6.5 Results and discussion 109
6.5.1 CG potentials 109
6.5.2 Representability of the CG models 112
6.6 Conclusions 114
6.7 Bibliography 115

7 Outlook 117

Summary

Molecular dynamic simulations provide valuable tools to study soft matter systems, since they can in principle shed light on many mechanisms happening at a size-scale hardly reachable experimentally. Unfortunately, despite the continuous development in computer technology and simulation algorithms, many phenomena characterizing soft matter still happen at a size and time scale that is not easily reachable by simulation carried with atomistic resolution. Coarse graining (CG) the system by representing groups of atoms as single interaction sites (beads) enables the possibility of studying such phenomena. The quality of a coarse-grained model is related with the quality of the description of the interaction between the beads, which should properly account for the chemical nature of the groups of atoms that they represent. In the recent years several methods have been developed to obtain coarse-grained interaction potentials. In between the available approaches to CG of particular interest is the class of methods that undergoes under the name of systematic coarse graining. Common characteristic of the methods belonging to this class is to develop models using informations obtained at a more detailed level of description of the system (e.g. atomistic level). The quality of a CG model is normally evaluated discussing its ability to predict property at the state point considered during its development (representability) and also at different state points (transferability).

This thesis reviews the current state of the art of available systematic coarse graining methodologies and proposes the conditional reversible work (CRW) method as a new approach to develop interaction potentials. CRW is based on the calculation of the interaction free energy between the groups of atoms that will be represented by the beads, under the condition that they are embedded in their respective molecules. Unlike many other methods CRW delivers pair interaction potentials that are free from many-body contributions and that have a clear physical meaning. In principle different routes can be used to calculate the interaction free energy between the beads. In this thesis two different approaches are presented: the first one is based on a thermodynamic cycle that makes use of reversible work calculations and the second one is based on thermodynamic perturbation theory. The representability

and the transferability of coarse grained models obtained employing CRW methods are discussed. In particular first a comparison is carried between the performances of two toluene CG model developed employing CRW and a well-established CG method, namely Iterative Boltzmann Inversion (IBI), showing a remarkable transferability of the CRW model. Then a model of hexane where two neighboring carbons and their belonging hydrogens are lumped together in a single interaction site is developed using CRW methods. A comparison between the property of the parent atomistic model and of the CG model proves a good representability and transferability of the model. Considering the fact that the interaction potentials are representative of the conditional free energy of the group of atoms, also the chemical transferability is investigated. This expresses the possibility of employing the developed CG models for small molecules as building blocks for bigger molecules. The interaction potentials developed for the hexane molecules have been employed to simulate linear alkanes up to dodecane finding a good agreement between the properties calculated using the CG model and the parent atomistic model. Also the possibility of using CRW potentials for mixture has been tested investigating the behaviors of n-alkane mixture and the thermodynamic of solvating additives molecule in a polymeric matrix. Again a comparison between atomistic and CG results agree over a wide range of temperatures even for a sensitive quantities like the solvation free energy. In order to understand the limit of applicability of CRW pair potential approach also a study has been carried on the applicability of this method on systems that are governed by multi body interactions. In particular a series of increasingly polar liquid has been CG as single bead. It has been found that when directional interactions are CG away the model rapidly loses the ability of describing property of the system. It is important to notice that in principle it is possible to repair at least part of this failure by employing a lower level of CG (i.e. using more bead to describe the molecules), since this allows maintaining part of the directionality of the interactions.

This thesis proves the quality of CRW method to develop interaction potentials for soft matter systems. In particular it has to be remembered that the potentials developed using this method are pair potentials, free of indirect contributions. Therefore this method is straightforwardly applicable to develop coarse grained models for systems where pair additivity of the interaction can be considered like van der Waals dominated systems. A more careful implementation is necessary in system where multibody interactions are important. Since the CRW potentials are representative

of a conditional free energy it is possible to think future application of this methods also to develop models for systems where this kind of interaction are important, like hydrophobic interaction between amino acids in water.

Zusammenfassung

Molekulardynamische Simulationen leisten ein wichtigen Beitrag zum Studium der Weichen Materie, da sie Mechanismen auf mikroskopischer Ebene illustrieren können, die experimentell kaum zugänglich wären. Leider finden viele charakteristische Phänomene der Weichen Materie auf Längen- und Zeitskalen statt, die trotz der kontinuierlichen Weiterentwicklungen der Computertechnologie und Simulationsalgorithmen für atomistische Simulationen schwer erreichbar bleiben. Mit sogenannten vergröberten Modellen, coarse grained (CG) Modellen, bei denen mehrere Atome zu einem Superatom (CG bead) zusammengefasst werden, können ebendiese Längen- und Zeitskalen erreicht werden. Die Eigenschaften bzw. die Qualität dieser CG Modelle ist abhängig von der Beschreibung der Wechselwirkung zwischen den CG beads, wobei Letztere die chemische Natur der zugrundeliegenden atomistischen Details widerspiegeln sollten. In den vergangen Jahren wurden mehrere Methoden entwickelt, um die CG Wechselwirkungen zwischen einzelnen CG beads zu beschreiben, dazu zählt auch das sogenannte systematische coarse graining. Hierbei werden Information von der atomistischen, also chemisch detaillierteren, Ebene für die Entwicklung dieser CG Modelle herangezogen. Die Qualität dieser CG Modelle wird häufig daran gemessen, inwiefern diese Modelle, die bei einer bestimmten Temperatur und Druck entwickelt wurden, andere physikalische Eigenschaften bei gleichen Bedingungen (representability) oder anderen Bedingungen (transferability) vorhersagen können. Diese Arbeit fasst den derzeitigen Wissensstand auf dem Gebiet des systematischen coarse graining Methoden zusammen und führt einen neuartigen Ansatz, die sogenannte conditional reversible work (CRW) Methode, zur Entwicklung von CG Wechselwirkungspotentialen ein. CRW basiert auf der Berechnung der freien Wechselwirkungsenergie zwischen Atomgruppen, die später die CG beads darstellen, unter der Voraussetzung, dass sie während der Berechnung in dem Molekül eingebettet sind. Im Gegensatz zu anderen CG Methoden, sind die CRW Wechselwirkungspotentiale frei von sogenannten multibody contributions und besitzen eine eindeutige physikalische Bedeutung. Prinzipiell können diese CG Wechselwirkungspotentiale auf mehreren Wegen berechnet werden. In dieser Arbeit werden zwei dieser Wege diskutiert: Der erste basiert auf der Berechnung von eines thermodynamischen

Kreisprozesses und der zweite basiert auf der thermodynamischen Störungstheorie. In dieser Arbeit werden representability und transferability der CRW CG Modelle diskutiert. Insbesondere der Vergleich eines CRW Modells für Toluol mit einem Modell, welches auf mit der sogenannten "Iterative Boltzmann Inversion" (IBI) Methode entwickelt wurde, zeigt, dass die CRW Wechselwirkungspotentiale eine bemerkenswerte transferability besitzen. Diese Ergebnisse werden weiterhin durch die Studie eines CRW CG Modells für Hexan unterstützt und durch einen Vergleich mit dem atomistischen Modell, auf dem das CG Modell basiert. Des Weiteren wird in dieser Arbeit die chemische transferability der CRW CG Modelle analysiert. Dies ermöglicht, dass die CG CRW Wechselwirkungspotentials für kleine Moleküle in einer Art Baukastensystem für größere Moleküle verwendet werden können. Im Speziellen wird die chemische transferability der CRW Wechselwirkungspotential für Hexan anhand von längeren, linearen Alkane (bis zu Dodekan) getestet. Hierbei zeigen die CG Berechnungen gute Übereinstimmungen mit den atomistischen Berechnungen. Des Weiteren wird in dieser Arbeit CRW Modelle für Mischungen anhand zweier Beispiele diskutiert: Mischungen von n-Alkanen und die Solvatisierung von Additiven in Polymeren. Auch hier zeigen sich, sogar für empfindliche Größen, wie z.B. die freie Solvatisierungsenergie, gute Übereinstimmungen von CG und atomistischem Modell für einen weiten Temperaturbereich. Um die Grenzen der Anwendbarkeit des CRW Ansatzes zu verstehen, wurde die Anwendbarkeit dieser Methode für Systeme untersucht, die von multibody Wechselwirkungen dominiert werden. Genauer wurden Flüssigkeiten steigender Polarität durch eine einzige CG bead dargestellt. Es zeigte sich, dass die Modelle die Eigenschaften des Systems nicht mehr beschreiben können, sobald richtungsabhängige Wechselwirkungen durch die Vergröberung verloren gehen. Allerdings ist es prinzipiell möglich diesen Fehler zu beheben, indem man eine weniger vergröberte Darstellung wählt (also mehr CG beads zur Beschreibung eines Moleküls verwendet), da dies die Richtungsabhängigkeit der Wechselwirkungen teilweise erhalten kann.

Diese Arbeit belegt die Qualität der CRW Methode zur Entwicklung von Wechselwirkungspotentialen für Systeme Weicher Materie. Besonders hervorzuheben ist, dass es sich bei den entwickelten Potentialen um Paarpotentiale handelt, die frei von indirekten Beiträgen sind. Deshalb ist diese Methode unkompliziert zur Entwicklung von CG Modellen für Systeme anwendbar, in denen die paarweise Additivität der Wechselwirkungen wie in van-der-Waals-Systeme angenommen werden kann.

Systeme mit erheblichen multibody Wechselwirkungen erfordern mehr Sorgfalt bei der Umsetzung. Da die CRW Potentiale eine bedingte freie Energie repräsentieren, wäre es möglich sich eine Anwendung dieser Methoden zur Entwicklung von Modellen für Systeme vorzustellen, in denen diese Wechselwirkungen wichtig sind, wie hydrophobe Wechselwirkungen zwischen Aminosäuren und Wasser.

Thanks!

I'll use this page to take the pleasure to acknowledge the many people who made the path trough my PhD years an unforgettable experience.
First and foremost, my gratitude goes to Nico van der Vegt. With his overwhelming entusiasm for science and life he deeply influenced me. During this three years he allowed me to follow my scientific curiosity and inspirations, providing always an invaluable support made of good teaching, good company, good advices and a lot of great ideas. I honestly can not immagine a supervisor better suited for me. I wish to express my sincere gratitude also to Prof. Müller-Plathe for all what I learned during the *theory group seminars* and his lectures. A special thanks goes to Prof. Dr. Robert Berger and Prof. Annette Brunsen that accepted to be part of my PhD committee, allowing the final discussion of my thesis.
It is also time to thank the many friends with whom I shared the last three years: without them for sure I will not be here today. My gratitude goes to all my working colleagues. Together we had plenty of good time (during and outside working hours) and they taught me many things both about science and life. So thank you Valentina, Claudia, Chunli, Pim, Simon, Enrico, Davide, Gregor, David, Pritam, Timir, Fereshte and all my other beloved colleagues. For sure my experience in Germany would have not been the same if I would not have joined the Darmstadt Rugby team. Thank you all guys for the unforgettable time we spent together both on and off the pitch! I also have a debt of gratitude with my friends in Italy since every time I went back it was like I had never left. Thanks Bito, Sara, Frank, Vale, Ale, Rino, Paride, Mona, Max, Luca, Francesca e Icio.
Last but not least I thank my parents for their continuous support and the love I constantly received in these 28 years.

Emiliano

1 Introduction

Understanding phenomena happening at the nanoscale is crucial for the study of soft matter; simulations provide valuable tools to investigate them. Despite the tremendous improvement in algorithms and in computer power during the last few decades, simulating soft matter systems with full atomistic resolution is still in most cases a prohibitive task. In order to overcome this limitation it is possible to employ simplified *coarse grained* (CG) models where groups of atoms are lumped together in *beads*. To preserve an accurate description of the system it is necessary that the interactions between the beads are representative of the chemical nature of the groups of atoms that they represent. Two different approaches can be employed to build CG models: in the first case the system is parametrized in order to reproduce macroscopic properties[1–3] in a similar way to what it is done for the parametrization of many atomistic force fields; in the second case CG models are built using information acquired at the atomistic level of resolution. This second kind of approach is usually identified as *systematic CG*. Systematic CG methods can be further subdivided between the ones that parametrize models in order to reproduce microscopic properties[4–9] and the ones that uses interactions calculated between groups of atoms at the atomistic level as interaction between the CG beads[10–12]. The first approach to systematic CG ensures to obtain models able to reproduce target properties, but nothing is a priori said about the ability of these models to reproduce any other property at the same or at a different state point from the one considered during the development of the model. On the other hand nothing is said about the ability of the models developed with the second groups of methods to reproduce any property; nevertheless when they are able to capture the nature of the interactions of the system the fact that they deliver potentials with a clear link with a physical interaction makes them more likely to reproduce several property of the system. Even more the clear link of these potentials with physical quantities enables route to understand and repair eventual failure of the models.

In this thesis we introduce a new CG method belonging to the second class of approaches to systematic CG: the conditional reversible work (CRW) method. In particular we define the interactions between the beads as the interaction free energy

between the groups of atoms that the beads represent calculated under the condition that they are embedded in their respective molecules. The choice to use a free energy to describe the interactions between the beads comes naturally considering the differences between an atomistic and a CG representation of the system. In fact in CG representations the number of interaction sites is smaller than in atomistic ones: this means that a part of the configurational entropy of the systems is lost. Therefore a potential should account both for the energetic interaction between the groups of atoms and for the loss of configurational entropy: free energy naturally includes these two contributions. It is interesting to note that the calculated free energy between the groups of atoms is a conditional free energy, this means that the interaction between the beads is representative of the configurational space sampled by the two groups of atoms, that it is influenced by the chemical nature of their environment. An other key property of CRW derived potentials is that, unlikely in many other approaches[4–10], they represent a pair interaction free of indirect contributions. This is a clear advantage for the simulations of systems dominated by pairwise additive interactions (e.g. apolar liquid), but this also introduces limitation in representing systems where multibody interactions are important (e.g. water).

The thesis is structured as follows. Chapter 2 reviews the current available systematic CG methods, and summarizes pro and con of the different approaches to systematic CG in relation to four different system classes: non polar liquids, water and polar liquids, polymers and in particular polystyrene, and ionic solutions. These four systems are dominated by different molecular interactions, and it is interesting to understand how different systematic CG methods perform in "reproducing" them. In chapter 3 the CRW method is presented and in particular a route based on the use of a thermodynamic cycle is employed to calculate the interaction free energy between the groups of atoms. This strategy is adopted to derive three beads models for toluene and for hexane. These show remarkable ability to predict thermodynamic properties over a wide range of temperature. In chapter 3 it is also shown that it is possible to carry the sampling of the interaction between the two groups of atoms in vacuum or in bulk liquid. In this second case the molecules during the sampling of the interactions explore the phase space of the bulk liquid, and this improves the ability of the model to reproduce properties associated with this state. In chapter 4 a new method based on thermodynamic perturbation theory is presented in order to calculate CRW interaction potentials. In the same chapter also the importance of

the use of a conditional free energy is highlighted showing the possibility of employing the interaction potentials developed for a 3 beads model of hexane as building block for longer linear alkane chains. It is shown that properties predicted by the CG models and by the parent atomistic one are in good agreement over a wide range of temperature for linear alkanes with an even number of carbon atoms from butane to dodecane. In this chapter are also investigated properties of linear alkanes mixtures both at the atomistic and at the CG level, finding a satisfactory agreement with atomistic and experimental results. In chapter 5 the solvation thermodynamics of additive in polymer is investigated using CG models. It is shown that the CRW developed CG models are able to predict with a reasonable agreement the excess free energy associated with the solvation of ethylbenzene, methane and neopentane in a polystyrene matrix. Also the discrepancies between the predictions of thermodynamic properties of CG and parent atomistic models are carefully investigated. CRW method, unlikely many other CG methods, delivers pair interaction potentials free of multibody contributions, therefore CRW developed CG models are easily applicable in systems characterized by this kind of interactions. In chapter 6 the possibility of employing CRW for systems where multi-body interactions are important is investigated. In particular single bead models are built for bulk liquid toluene, dimethil ether, acetone and dimethil sulfoxide. In this chapter it is shown that the assumption of pair additivity of the interaction potentials is pretty strong in these systems since no one of the models is really representative of its target system. This sets a clear limit to the applicability of CRW developed models. Nevertheless it is interesting to understand that CRW models of polar molecules can in principle be built employing multi-beads representation in order to keep a certain certain level of directionality of the interactions. Finally in the outlook chapter results are summarized and possible future applications of the CRW method are discussed.

1.1 Bibliography

[1] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *J. Phys. Chem. B*, 111:7812, 2007.

[2] K. A. Maerzke and J. I. Siepmann. *J. Phys. Chem. B*, 115:3452, 2011.

[3] J. R. Allison, S. Riniker, and W. F. van Gunsteren. *J. Chem. Phys.*, 136:054505, 2012.

[4] F. Mueller-Plathe. *Chem. Phys. Chem.*, 3:754, 2002.

[5] A. Lyubartsev and A. Laaksonen. *Phys. Rev. E*, 52:3730, 1995.

[6] M. S. Shell. *J. Chem. Phys.*, 129:144108, 2008.

[7] F. Ercolessi and J. B. Adams. *Europhys. Lett.*, 26:583, 1994.

[8] S. Izvekov and G. A. Voth. *J. Phys. Chem. B*, 109:2469, 2005.

[9] J. W. Mullinax and W. G. Noid. *J. Phys. Chem. C*, 114:5661, 2010.

[10] B. Hess, C. Holm, and N. F. A. van der Vegt. *J. Chem. Phys.*, 124:164509, 2006.

[11] Y. Wang, W. G. Noid, P. Liu, and G. A. Voth. *Phys. Chem. Chem. Phys.*, 11:2002, 2009.

[12] E. Brini, V. Marcon, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 13:10468, 2011.

Publications

This thesis is based on the following publications:

E. Brini, E. Algaer, C. Li, P. Ganguly, F. Rodrìguez-Ropero, and N. F. A. van der Vegt, *Systematic coarse graining methods for soft matter simulation: a review*, Soft Matter, (submitted)

E. Brini, V. Marcon, and N. F. A. van der Vegt, *Conditional reversible work method for molecular coarse graining applications*, Phys. Chem. Chem. Phys., 13:10468 (2011)
Reproduced by permission of the PCCP Owner Societies

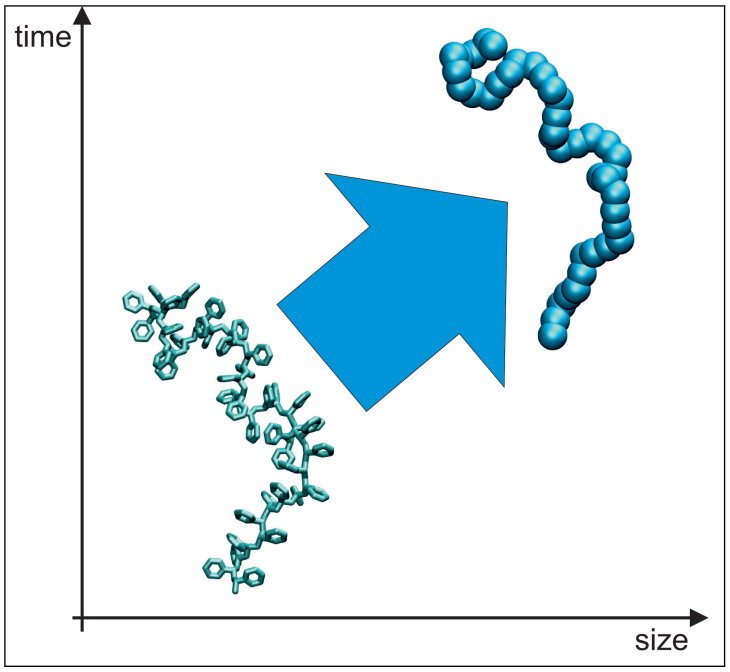
E. Brini and N. F. A. van der Vegt, *Chemically transferable coarse-grained potentials from conditional reversible work calculations*, J. Chem. Phys., (submitted)

E. Brini*, C. R. Herbers*, G. Deichmann, and N. F. A. van der Vegt, *Thermodynamic transferability of coarse grained potentials for polymer-additive systems*, Phys. Chem. Chem. Phys., 14:11896 (2012)
Reproduced by permission of the PCCP Owner Societies

* shared authorship

2 Systematic coarse graining methods for soft matter simualtion: a review

Multiscale simulations of soft matter is an emerging field that has made rapid progress in the past decade. Several methods for systematic coarse-graining of molecular liquids and soft matter systems have emerged in recent years. Herein, we review these methods, discuss a selected number of applications as well as limitations of the models and remaining challenges in developing representative and transferable pair potentials.



2.1 Introduction

The properties of soft condensed matter are determined by structure and dynamics in an extremely wide range of time and length scales. Theoretical modelling descriptions of soft matter systems are therefore confronted with the challenge of scale bridging in order to link chemical structure and properties. Systematic (bottom up) coarse graining approaches have been very successful in recent years in addressing physical questions that require a certain amount of chemical specificity being represented in the models. Typical questions of this kind involve modelling of, *e.g.*, protein interactions with lipid bilayers, self-assembly processes, polymer material and surface properties, and so on. Classical molecular dynamics (MD) simulations with all-atom force field models are limited to very small systems and nanosecond time scales, therefore the development of simplified or *coarse-grained* (CG) molecular models by systematic coarse graining has become an active field of research in the past decade. Coarse graining molecules by representing groups of atoms as single interaction sites facilitates the study of soft matter systems since (i) the total number of particles present in the system is reduced, (ii) the interaction potentials are simplified (iii) and the potential energy surface on which the molecules move is smoothed, leading to an acceleration of the molecular dynamics.[1]

In almost all coarse graining procedures effective pair potentials are derived which are computationally efficient but restricted in their ability to represent all system properties. Potentials can be developed in order to reproduce mainly microscopic quantities obtained from fine-grained (*e.g.* detailed-atomistic) simulations[2–10] (*bottom up approach*) or macroscopic -thermodynamic- quantities[11–13] in a similar fashion to the parametrization of most of the atomistic force-fields[14–16] (*top down approach*). The line that separates these two methodological approaches is actually extremely thin as it is demonstrated from the fact that both approaches can be combined to develop coarse grained potentials.[12] In order to define the range of applicability of a developed model, it is necessary to test the ability of the model to predict properties at the thermodynamic state point used during its development (*representability*) and at different state points (*transferability*).

In this work we review a series of bottom up coarse graining methods, comparing the performance of different CG models developed using different methods on similar systems. We focus our attention on four different system classes: apolar molecular

liquids, water and polar molecular liquids, polymers and in particular polystyrene (PS), and ionic solutions.

2.2 CG interaction potentials

In the process of coarse graining a series of relevant, slow degrees of freedom (DoFs) is defined by a mapping scheme. The mapping usually involves representing a small number of atoms by a single interaction site, and depends on the amount of detail the CG model should retain in order to investigate the physical problem of interest. The CG DoFs ("slow" DoFs) are denoted $\vec{R}$, while the DoFs that are lost by performing the mapping ("fast" DoFs) are denoted $\vec{r}$. The partition function Q of a system then reads

$$Q = \int d\vec{R} \int d\vec{r}\, e^{-\beta U^{AA}(\vec{r},\vec{R})} \tag{2.1}$$

where $\beta = (k_B T)^{-1}$ with k_B the Boltzmann constant and T the absolute temperature, and U^{AA} is the fine-grained (all-atom) interaction potential. The multibody potential of mean force $U_{eff}(\vec{R})$, which governs the dynamics of the coarse-grained system $\vec{R}$, is defined as

$$U_{eff}(\vec{R}) = -k_B T \ln \int d\vec{r}\, e^{-\beta U^{AA}(\vec{r};\vec{R})} \tag{2.2}$$

and represents the free energy of the lost degrees of freedom for CG configuration $\vec{R}$. With this definition it follows that

$$Q = \int d\vec{R}\, e^{-\beta U_{eff}(\vec{R})} \tag{2.3}$$

Although this "ab-initio" coarse graining procedure is thermodynamically consistent, the effective potential (eq.2.2) is not practical in computer simulations due to its multibody nature. In all coarse graining methods reviewed below, effective *pair* potentials are used instead.

The transferability and representability of a CG model are limited by two factors: the nature of the interaction potentials that is not purely energetic and the fact that

for computational reasons the interaction potentials are mainly implemented as pair potentials. The first limitation comes from the coarse graining process itself. Averaging over fast DoFs removes their explicit entropic contributions. To compensate for this, the pair potentials usually exhibit somewhat weaker attractions and repulsions resulting in a potential energy surface that is flattened out in comparison to the fine-grained atomistic system (see Fig. 2.1). Fluctuations resulting from sampling of the CG energy surface are different from the fluctuations in the atomistic system, resulting in different thermodynamic response functions, and, therefore, limited state-point transferability of the CG model. For example, the heat capacity, which, at constant pressure and temperature, is related to enthalpy fluctuations, will be underestimated by the CG model while the compressibility, which is related to volume fluctuations, will be overestimated. The thermal expansion coefficient, which is related to correlated volume and enthalpy fluctuations, may however still be represented correctly. In many systematically coarse-grained models the number of atoms represented by a CG bead is relatively small, and a certain grade of transferability is often achieved across a finite range of temperatures owing to an only moderate flattening of the potential energy surface. The second limitation to the representability and the transferability of the CG model comes from practical reasons of implementing the CG model and keeping the simulation inexpensive. To achieve this, the interaction potentials are usually assumed to be pairwise additive. This assumption, which may be reasonable for specific classes of compounds like van der Waals fluids, however limits the representability of the CG model.

Interaction potentials between the CG beads of a molecular system usually distinguish between bonded and non-bonded contributions. In the next two sections, we discuss the different methods available to develop both of these contributions.

2.3 CG bonded potentials

CG bonded potentials are used to model the flexibility of degrees of freedom associated with coarse-grained bonds, angles and dihedrals. A straightforward way to obtain these potentials is by Boltzmann-inverting the sampled distribution of the mapping points from a reference fine-grained simulation.[17] Since most bonded DoFs can usually be considered "stiff" (note that this may not be the case for the dihedrals), no strong coupling is present with the weaker nonbonded interactions. This justifies the direct Boltzmann inversion of sampled bonded distributions and the assumption of additivity of the CG bonded and nonbonded interactions. A further

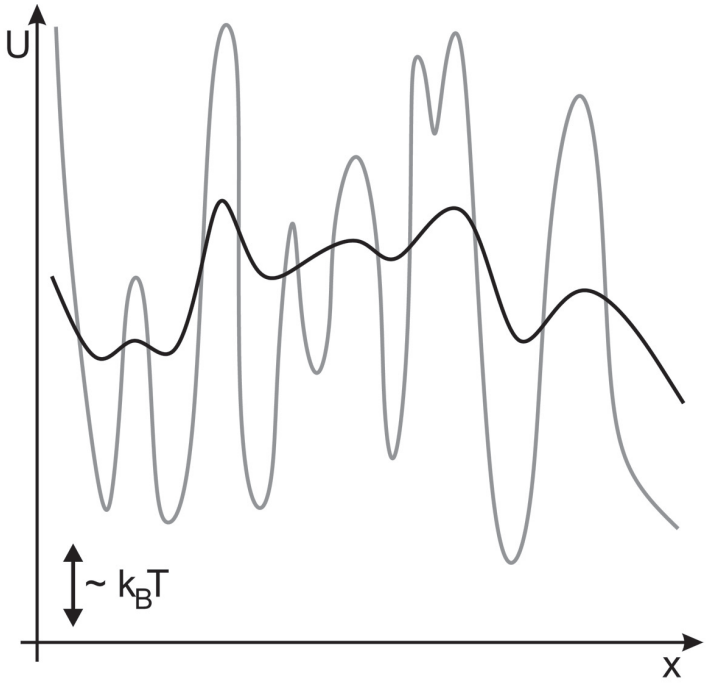


Figure 2.1: Schematic representation of a 1D potential energy surface for a hypothetic atomistic system (gray line) and its CG representation (black line). The CG surface is smoother than the atomistic surface, therefore the CG system samples its configurational space faster. The transferability of the CG model is however limited because the equilibrium fluctuations that determine the thermodynamic response to changes in temperature and pressure are different at the two resolution levels.

assumption which is usually made relies on the motions of CG bonds, angles and dihedrals to be independent, and the CG bonded potentials to be modelled as additive sums of these DoFs. The validity of this assumption, however, depends strongly on the chosen mapping scheme and can not be guaranteed in general. [18] In some cases, long-range bonded potentials (*e.g.* 1-5 bonded potentials) can be introduced to reproduce the local conformational characteristics of flexible molecules in agreement with the fine-grained model as was demonstrated by Fritz *et al.* [18]

In case bonded and nonbonded interactions are not independent, an alternative approach is needed. In this case, the bonded potential can be iteratively refined until distribution functions of bonded DoFs obtained from fine-grained simulations are reproduced [6,19]

2.4 CG nonbonded potentials

Nonbonded interactions determine many of the properties of soft matter systems. It is therefore crucial to develop methods able to derive these interactions for coarse-grained systems from atomistic models. Two different methodological approaches can be identified in the literature. In the first approach, which we define as *parametrized* coarse graining, atomistic simulations are used to calculate one or more target properties (*e.g.* pair correlations or force distribution), followed by deriving CG potentials, which, by construction, reproduce these target quantities. It cannot be guaranteed that these models can predict other, non-target properties at the same or at a different state point from the one used in the parametrization. The second class of methods (*derived* coarse graining) uses the direct atomistic interactions between the mapped atom groups to derive the CG interaction potentials. CG potentials belonging to this class are not optimized to reproduce preselected targets, hence all calculated properties are predictions of the model. These potentials have a clear physical meaning, representing a distance-dependent pair interaction (free) energy and, as will be shown below, exhibit a good transferability. Derived pair potentials can be systematically modified to include effective contributions, *e.g.*, to correct for multibody solvent effects in implicit-solvent models.[8,20,21] A scheme of the available systematic coarse graining methods is shown in Fig. 2.2.

2.4.1 Parametrized methods

Parametrized methods are those methods that provide CG models able to reproduce by construction a target radial distribution function (RDF) or a target force distribution calculated from simulations carried out at a more detailed level of resolution (*e.g.* atomistic simulations).

All methods that aim at the reproduction of a target RDF (*structure based* coarse graining) rely on the Henderson uniqueness theorem, which states that there is only one pair potential able to exactly reproduce a given RDF.[23] However it has been shown that in practice a series of potentials are normally able to reproduce a given structure within an acceptable error. Due to the straightforward implementation and the robustness of its algorithm, one of the most used structure-based CG procedures is the **Iterative Boltzmann Inversion** (IBI) method, proposed by Reith *et al.*[2,19,24] This method aims at the construction of a tabulated potential able to re-

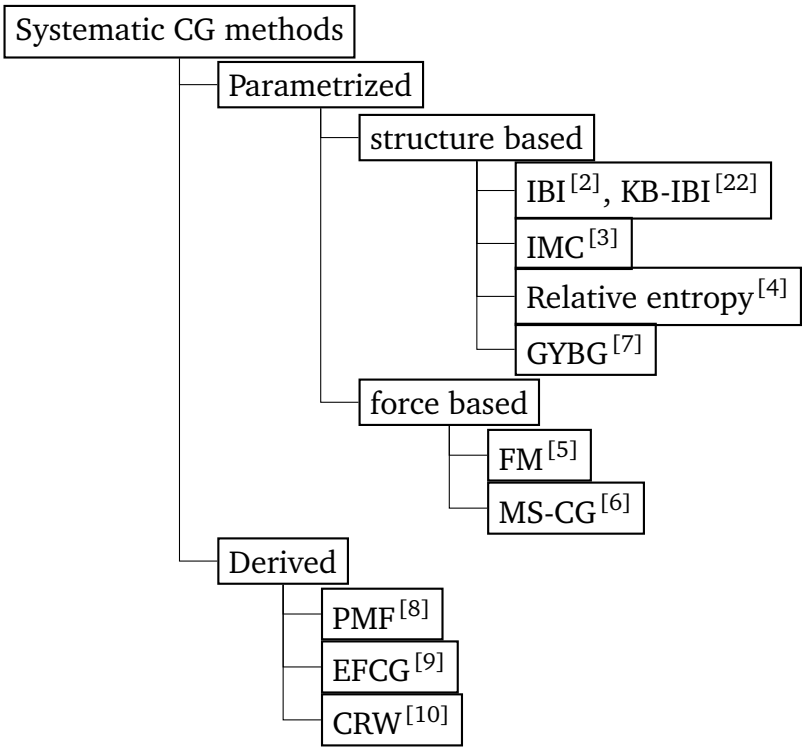


Figure 2.2: General scheme of the available systematic CG methods

produce a target radial distribution function $g(r)$ from atomistic simulations. The approach is similar in philosophy to an earlier developed methodology used to iteratively derive atomistic pair potentials able to reproduce structure factors determined from neutron scattering experiments.[25–27] In order to derive the interaction potential, the IBI method follows a procedure where an initial guessed CG pair potential is iteratively refined using a correction proportional to the difference between the fine-grained (atomistic) and coarse-grained potentials of mean force $-k_B T \ln g(r)$. The correction is made assuming one-to-one correspondence between the potential at a given distance r and the value of the RDF at the same distance.[2] Considering the fact that many pair potentials are able to reproduce a target $g(r)$ within an acceptable error, additional thermodynamic target properties can be used in the iterative scheme, like the pressure,[2] or in the case of multi-component mixtures, the Kirkwood Buff integrals, which relate to the activity coefficients of the species in solution.[22] The latter approach, coined the KB-IBI method, links structure and thermodynamics and has been recently applied to urea-water mixtures.[22] This link is particularly important in modelling biomolecular systems where variations in the activity of solvent components affect biomolecular stability. An alternative way to improve the quality of IBI potentials and the speed of convergence of the method is to rely on a high quality initial guess of the potential to refine. This is the case of the

so called Inverted Ornstein-Zernike equation with the Percus-Yervick approximation ($OZPY^{-1}$).[28]

Even though it was developed earlier, a conceptual evolution of the IBI method is the so called **Inverse Monte Carlo** (IMC) method first presented by Lyubartsev and Laaksonen.[3] IMC also relies on an iterative scheme that corrects a guessed interaction potential in order to reproduce a target radial distribution function, but the correction applied to the potential is based on statistical mechanical arguments that take into account that a variation of the potential at a given distance r can lead to a variation in the radial distribution function at all other distances. This relation can be treated analytically leading to a rigorous iterative scheme.[3] Based on that, the IMC method should lead to a faster convergence of the potential compared with IBI and in principle it will remove eventual convergence problems in multi-component mixtures due to correlations between the observables. This however comes at the price of computational cost since it is necessary to calculate correlations between updates in potentials and radial distribution functions at different distances, and in order to obtain a proper statistics to calculate those cross correlation long simulation runs are necessary.[29] The consistency of this method has been proven by employing IMC to re-obtain the original interaction potentials between particles that generated a particular pair structure in systems governed by spherically symmetric interactions.[30] Also in the case of IMC it is possible to include target thermodynamic properties like the surface tension into the iterative scheme.[31,32]

It can be shown that both IMC[33] and IBI[34] are particular cases of the **relative entropy** formalism introduced by Shell,[4] that provides a general framework based on the relative entropy to optimize a CG forcefield. The relative entropy is a positive definite measure of the amount of information lost upon coarse-graining and is defined in terms of probability distributions of the atomistic and coarse-grained systems. The CG model is obtained by minimizing the relative entropy, *i.e.* by minimizing the discrepancies between the atomistic and coarse-grained probability density distributions. The relative entropy method can be applied to optimize a broad range of parameters following an optimization procedure over a model parameter space, or alternatively, to iteratively refine a CG model until the measure of the relative entropy converges to a minimum.

The second group of parametrized CG methods is based on the idea of matching the force distributions on the beads in the CG models with the force distributions on the mapping points of the atomistic model. The **Force-Matching** (FM) method was

first reported by Ercolessi and Adams[5] for obtaining potentials by fitting atomic forces and trajectories from $ab\ initio$ calculations. The fitting rapidly becomes intractable as the number of parameters grows, the FM method was therefore further extended under the name of **Multi Scale Coarse Graining** (MSCG) by Izvekov and Voth.[6,35] The FM/MS-CG method derives *effective* pair potentials that provide the best approximation of the multibody potential of mean force. A rigorous statistical mechanical description of this method was given by Noid $et\ al.$[36,37] In principle, if the FM is carried out unrestricted, it provides the same results as the relative entropy method, but the use of *effective* pair potentials that are not able to directly capture multibody interactions leads the two methods to deliver different results.[34,38] The similarity between force and structure based method is at the base of the idea of the **generalized-Yvon-Born-Green** theory (GYBG) developed by Mullinax and Noid.[7] The authors show that a forcefield developed solving a set of a linear integral equations expressed in terms of structural correlation functions of the CG sites provides an optimal approximation of the multibody PMFs. Therefore, employing structural correlation functions of mapped atomistic configurations, it is possible to determine the projection of the multibody-PMF onto a CG (pairwise) force field; in other words it is possible to derive a CG pair potential that is the best approximation to the multibody-PMF directly from structural atomistic information.

2.4.2 Derived Methods

Derived methods are those in which the CG pair interaction potentials are calculated directly at the atomistic level from direct interactions between the groups of atoms represented by the beads. These methods thus provide pair interaction potentials, which can be derived for pure fluids but also for solution systems with an implicit or explicit representation of the solvent. The implicit-solvent CG pair potential should be considered an *effective* pair potential that includes the indirect (multibody) contributions of the removed solvent, while for the other cases it should be considered a pure pair potential. In this last case the CG procedure will deliver meaningful results only in systems where the pair interactions are dominant with respect to the other interactions (e.g. Lennard-Jones dominated system). We divide the method in this class accordingly to the property used to describe the interactions between the beads.

In solute-solvent systems the **pair Potential of Mean force** (pPMF), calculated in the limit of low solute concentration (two solutes in a solvent box), has been used

as an effective interaction in implicit solvent models. The pPMF includes indirect multibody contributions of the (removed) solvent and must therefore be considered an *effective* pair potential. Because the multibody contributions depend on solute concentration, a limited concentration range can be studied with these types of potentials. This concept was first studied by Hess *et al.* [8,21] for simulating implicit-solvent electrolyte solutions. There it was shown that transferability of the pPMF-based potentials to higer salt concentrations could be achieved by introducing a salt-dependent dielectric constant in the long-range Coulomb contribution to the pPMF-based potential. This methodology has more recently been further extended to bigger molecules like peptides [39] and polyelectrolytes. [40] In the case of bigger molecules the interactions between the CG beads have been calculated as interaction between chemical fragments. [39,40] For example, the interaction between two phenyl rings in a polystyrene sulfonate monomer is calculated as the interaction between two benzene molecules. [40] This implies the assumption of a certain degree of transferability of the interaction potential to a different chemical environment. The development of CG models using this approach is referred to as *fragment-based* coarse-graining.

Two kinds of methods have been reproted in the literature to derive CG pair potentials. The first one is the effective force, and the method goes under the name of **Effective Force CG** (EFCG). [9] This method explicitly computes the total translational force between two groups of atoms and project it onto the radial vector between the mapping point. At a given distance the forces are averaged over the different configurations allowed by the surrounding environment, that does not contribute to the pair potential in any other way. The effective-average force is computed during a simulation as a function of the distance and it is used as direct input in MD simulations.

The other method that delivers pair potential is the **Conditional Reversible Work** (CRW) method of Brini *et al.* [10] This method calculates the interaction free energy between the groups of atoms that the beads represent at all distances sampled in a fine-grained atomistic simulation and uses that quantity as a pair potential in the CG simulation. The CRW pair potential is calculated under the condition that the groups of atoms are embedded in the chemical environment defined by the molecules they belong to. The use of the free energy guarantees that the potential naturally includes the interaction energy of the atoms in the two groups (weighted over the relative configurations) and the configurational entropy of the two groups of atoms that is

lost during the CG process. The fact that the interaction free energy is calculated under the condition that the group of atoms are embedded in their respective molecules accounts for the proper weight of the relative orientation, since only the orientation allowed by the presence of the other atoms in the molecules are sampled. The sampling of the interaction free energy can be carried out between two molecules in vacuum or in liquid phase. In this second case the molecules sample the relative orientation allowed by the presence of the surrounding fluid.[10]

2.5 Applications

As previously stated, in this paper we will discuss the performance of different CG methods applied to different systems, trying to compare the difference in transferability and representability of the models developed with different methods.

2.5.1 Apolar molecular liquids

The first class of compounds that we investigate are apolar (or slightly polar) molecular liquids. Fig.2.3 show the mapping schemes of the systems which we will discuss. In this class of compounds the interactions between atoms are dominated by van der Waals forces, therefore it is reasonable to assume pairwise additivity of the CG potentials. It should be noted that for these types of systems composed of small molecules the thermodynamic properties are mostly determined by the non-bonded interactions. For these two reasons it is interesting to compare different coarse graining methods for this class of compounds.

Within the so called **parametrized CG methods** Ruehle *et al.*[29] compared the performances of IBI and FM methods by developing a united atom like force field for propane starting from an all-atom representation. They showed that the two methods lead to similar interaction potentials, and both models are able to reproduce the atomistic pair correlations calculated between the mapping points, showing therefore the ability of the FM methods to predict properties not accounted for during the parametrization. Similar conclusions were obtained for liquid hexane a in follow up work of Ruehle and Junghans.[45] Several studies have been published on the transferability of IBI-parametrized models for apolar solvents. Qian *et al.*[41] studied the transferability of IBI potentials for ethylbenzene at temperatures $T \neq T_0$, where T_0

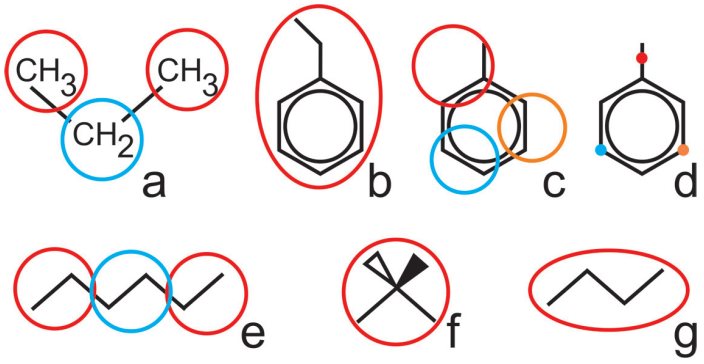


Figure 2.3: Mapping scheme of apolar molecular liquids considered in section 2.5.1. The CG methods employed include: (a) IBI[29] and FM[29]; (b) IBI[41]; (c) IBI[10] and CRW[10]; (d) GYBG[42]; (e) IBI[43] and CRW[10]; (f) EFCG[9]; (g) IBI[44], FM[44] and CRW.[44]

is the temperature at which the coarse-grained potential was derived. Application of such potentials at different temperatures evidenced a strong deviation of the thermal expansion coefficient predicted by the CG model both from the experimental value and atomistic molecular dynamics simulations. The same limitations were are also recently pointed out for liquid toluene.[10] To overcome this limitation the authors proposed the use of a scaling factor $f(T)$ to extrapolate such potential at different temperatures. A scaling factor $f(T) = (T/T_0)^{1/2}$ on the potential derived at 298 K allowed obtaining accurate potentials within the range 238K $< T <$ 380 K.[41] However the success of this methodology appears to be strongly dependent on the chosen T_0 and the analytical form of the $f(T)$. To overcome this disadvantage a more general approach to scale IBI derived potentials has been recently proposed by Farah *et al.*[43] This latter approach is based on a linear interpolation from the potentials obtained at two reference temperatures T_L and T_U, with $T_L < T < T_U$ and assuming that there is no phase transition in between. This strategy led to a very good fitting of the density and distribution functions in liquid hexane for temperatures ranging from 190 K to 338 K, but it is not straightforwardly applicable to nonequilibrium systems where a temperature gradient is present inside the simulation box (e.g. reverse non equilibrium MD calculations[46]). Employing the GYBG method Ellis *et al.*[42] developed two 3-site models for liquid toluene with slightly different mapping. Both models were able to reproduce the liquid packing of toluene in a range of 100 K. This proves the quality of this approach in system dominated by van der Waals forces.
As already mentioned in section 2.4, **derived methods** are expected to perform best for class of compounds where pair interactions are dominant. This is confirmed by the work of Brini *et al.*[10] who show the transferability of CRW-derived models of

toluene and hexane in a temperature range of about 100K. In the paper it is also shown that since the potentials have a clear physical meaning (i.e. interaction free energy) it is possible to understand and "repair" eventual failure of the model. In particular it is shown that for toluene the CRW method predicts a slightly lower excluded volume, which leads to an overestimation of the density. This can be repaired by slightly shifting the position of the beads in order to recover the right size of the molecule at least in the diameter of the ring. The EFCG method in this class of compounds has been applied to liquid neopentane. In particular the authors report the coarse graining of a neopentane molecule as a single interaction site, which is able to predict the pair structure of liquid neopentane. [9]

A comparison of the properties of CG models obtained from different methods is given in the work by Rzepiela *et al.* [44] There the authors compare the representability of CG models of single bead butane obtained with FM, IBI (identified as IB in the original paper) and CRW (identified as SB in the original paper) methods. They show that all the three methods are equally able to reproduce the pair structure of the liquid, but the CRW performs better in estimating the association constant as a function of the distance. Furthermore they show that in a simulation box at mixed resolution (atomistic-CG) the only model that does not show an appreciable preferential solvation between molecules at the same resolution is the CRW model. It has to be noted that at short distance a small preferential solvation between same resolution molecules is observed also with the CRW model, but this is almost lost already at the cutoff distance.

2.5.2 Polar molecular liquids - water

As opposite to the case presented in the previous section, in polar molecular liquids the multi-body contributions play an important role in determining the property of the systems and then they can not be neglected. In order to obtain simplified models where only pair interactions are considered, it is necessary that the interaction potentials are *effective* potentials carrying also information on multi-body interactions. These multi-body contributions are normally state point dependent, and this is often cause of a non optimal transferability of the CG models. The mapping schemes of the considered examples are reported in Fig.2.4

Almost all **parametrized** CG methods have been employed to obtain single-site models of water. The representability of these models is however limited due to

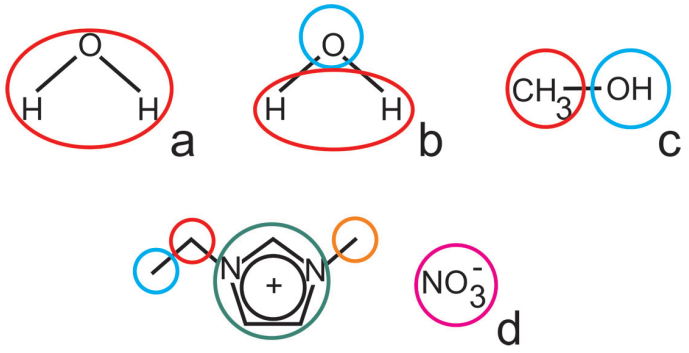


Figure 2.4: Mapping scheme of polar molecular liquids considered in section 2.5.2. The CG methods employed include: (a) IMC[47], IBI[48], MSCG[35,49,50] and relative entropy[4]; (b) MSCG[35]; (c) and (d) EFCG.[9]

important contributions of multibody interactions in this system. IMC has been employed to develop a CG single-site model for water,[47] starting from SPC water.[51] The aim of this model was to not only reproduce the structure of liquid water, but also its pressure at normal density (1 g/cm^3). IMC (or IBI) applied at normal density yields single-site models, which at the same temperature and density, have pressures of several thousand atmospheres. Lyubartsev *et al.*[47] however applied IMC to a phase separated system (a water droplet at equilibrium with its vapor phase at low pressure) and obtained a CG model that reproduces radial distribution function of bulk water (1 g/cm^3) as well as the pressure. This shows that by properly choosing the system to calculate a target property it is possible to obtain CG models that are able to reproduce not only the structure, but also other thermodynamic properties. It is interesting to notice that, in principle, it is possible to obtain a model that is able to reproduce structure and also thermodynamic property by setting up a constraint[32] or a correction[31] during the iterative procedure, but the approach used by Lyubartsev *et al.*[47] is more elegant since a thermodynamic property arises naturally from the iteration process. IBI coarse-grained potentials based on the TIP3P,[52] SPC[51] and SPC/E[53] non polarizable water models have been reported.[48] In that article each water molecule was described by a single bead. Authors obtained a good estimate of the compressibility factor compared with those from the reference all atom simulations when the pressure was not corrected, but this agreement was lost when the pressure correction was taken into account. The lifetime analysis of the tetrahedral cluster was in all cases slightly underestimated with respect to the value obtained from all atom simulations. This was improved by deepening the second minimum while increasing the height of the first peak of the water-water interaction potential (see Fig.2.5). In a more recent study, Hadley and McCabe[54] investigated

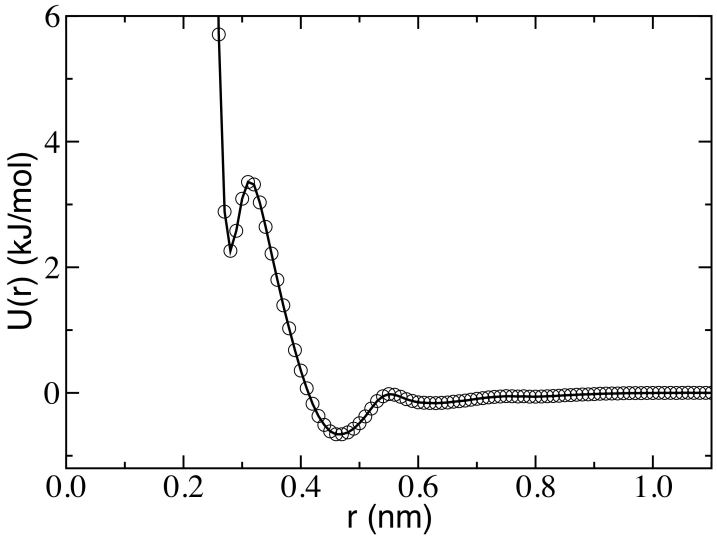


Figure 2.5: CG interaction potential for a single-site model of water obtained with IBI. In the work of Wang *et al.* [48] the author artificially changed the depth of the minima and the height of the peaks in order to obtain a CG model able to better reproduce properties of liquid water.

the optimal number of TIP3P water molecules mapped in one bead by using the K-means algorithm, [55,56] but this goes beyond the scope of this review.

In the work of Izvekov and Voth [35] three MSCG models for pure water were reported. The models were parametrized on the force distributions obtained from Car-Parrinello molecular-dynamics (CPMD). [57] The first model was a single-site model, the second model was a 2-site model without explicit charges and the third model was an explicit charges 2-site model able to reproduce the water dielectric constant. The first investigated property was the pair structure of water. The single-site model was able to reproduce it, while none of the 2-site models performed well. The authors justified this mismatch due to the difference in the hydrogen-bonded network in a 2-site model and a 3-site model as in all-atom H_2O. The thermodynamic quantities reported in the work, such as average configuration energy, heat capacity, thermal expansion coefficient and isothermal compressibilty are quite off for both 1-site and 2-site models compared to the reference atomistic simulation. Since the force matching was done using the virial constraint to obtain correct pressure, at 300 K all the models can reproduce the bulk density of water. Another 1-site MSCG water model was developed by the same authors [49] by fitting a reference simulation of rigid TIP3P [52] water molecules with an analytical polynomial for the short range force. This model was able to reproduce the CG-CG RDF of all-atom TIP3P water,

but failed to reproduce some thermodynamic properties.[52,58] In a later work[50] the necessity of a 3-body correlation function for water model was discussed. Using this 3-body correlation term a more accurate 1-site water model was developed by Larini *et al.*[59] The model, fitted to an atomistic simulation of SPC/E water,[53] was able to reproduce the water structure (radial distribution and angular distribution) more accurately than a 2-body potential. The relative entropy method has been employed to develop a spherically symmetric model of water based on a Lennard Jones potential superimposed to a gaussian.[4,60] The aim of those two works was not to develop a CG model, but to better understand hydrophobic interactions and their correlations with the water structure. Therefore a complete screening of water properties is not available, nevertheless even such a simple model can capture few properties of the hydrophobic interaction.

Not many studies have been reported based on **derived** models of polar liquids. Due to the pair nature of this potential a particular carefulness should be considered in order to include multibody contributions in the CG model. This can be done by employing partial charges on the beads after removing their effect from the CG interaction potential. Following this idea Wang *et al.*[9] developed a 2-site CG model for methanol using the EFCG method. In this model a bead represents the -OH group and the other bead represents the methyl group. In the paper they compared the radial distribution function of the CG model with the one of the parent UA model and found an acceptable agreement between the two RDFs. The authors compared also the surface tension of the two models. Considering this property is extremely sensitive to the interaction potential employed, the small discrepancies between the predicted values of the two models provides convincing proof of the quality of this coarse-graining approach. In the same paper the authors extend their idea to ionic liquids, showing a remarkable transferability of the obtained CG model to reproduce atomistic RDFs over a temperature range of 300 K. This shows that at least in principle pure pair potentials can be used to simulate systems where multi-body interactions are important at the computational price of leaving partial charges on the molecules.

2.5.3 Polymers - Polystyrene

Coarse graining of polymers is an extremely active field since the time needed to converge atomistic MD simulations of such systems is normally prohibitive. The most employed **parametrized** method to develop polymer models is IBI. We here

limit our discussion to polystyrene, which has been extensively studied with different coarse graining methods. IBI potentials for polystyrene have in all cases been derived using the all-atom force-field previously developed by Müller-Plathe,[61] in which Lennard-Jones parameters are taken from the benzene-benzene interaction of Jorgensen and Severance.[62] The mapping scheme shown in Fig. 2.6 a) was chosen to derive the first pressure-corrected coarse-grained potential via IBI for atactic PS.[63,64] This model, which was used in melts, reaching molecular weight up to 700.000, succeeded in reproducing the gyration radius and the Flory characteristic ratio at 500 K. However the entanglement length predicted by this model was underestimated with respect to the experimental value. Similar results were obtained by Sun and Faller[65,66] using a potential derived at 450 K without pressure correction (see Fig. 2.6b) where only isotactic polymer chains were considered. An optimum description of the entanglement length[67] was further obtained by slightly modifying the coarse-grained potential from the work of Milano *et al.*[63,64] in combination with the Contour Reduction Topological Analysis (CReTA) algorithm.[68] Additionally, other structural parameters such as the packing length or the reptation tube diameter were well reproduced compared with experiments.[69] However, the isothermal compressibility was largely overestimated, which shows the poor transferability of the obtained potential to pressures different from that used to derive the potential. Changing the mapping scheme (see Fig.2.6c), Qian *et al.* were able to reproduce also the isothermal compressibility and structural properties in the range from 400 K to 500 K.[41] IBI has also been successfully applied to develop the non bonding potentials of diblock copolymers such as poly(styrene-b-butadiene), where styrene and butadiene units are mapped using two different beads, both centered in the center of mass of each super-atom (see Fig. 2.6e).[70] In such case the three RDFs between the center of mass of the two beads must be adjusted to determine the three non-bonded potentials. Coarse-grained simulations using these potentials reproduced in a quite good agreement both static (gyration radius) and dynamic (diffusion coefficient) properties compared with all atom simulations.

Within the class of **derived** CG methods, Fritz *et al.*[18] developed a CG model of PS employing the CRW method. The mapping scheme of this model can be found in Fig.2.6d. The nonbonded potential were developed between two trimer chains in vacuum; it has been found that the nonbonded potentials calculated for syndiotactic sequences were identical to those of isotactic. The bonded potential were developed by Boltzmann inverting the probability distribution of CG bonds, angles and

dihedrals. To properly reproduce the stiffness of the polymeric chain a 1-5 bonded potential was also added. After developing the potentials in vacuum, the authors simulated polystyrene melts under isothermal-isobaric conditions at pressure of 1 atm and temperature of 503 K in order to compare atomistic and coarse-grained simulations. Chain conformations in coarse-grained and atomistic melt simulations have been compared using the characteristic ratio, showing that the model is able to predict correctly the different stiffnesses of PS chains with different tacticities. The local packing of polystyrene melt in the CG simulation is also reproduced in good agreement. To test the temperature transferability of the developed CG model, the authors performed a series of simulations in the range of temperatures between 403 and 523 K, and compared atomistic and CG densities. They found an excellent agreement between the two values. Moreover, the CG model describes the melt packing and reproduces the density of PS melt very well in the range of temperatures from 400 to 520 K, showing the transferability of such model. The approach chosen for the development of nonbonded interaction is computationally inexpensive since it is based on the atomistic simulation of trimer in vacuum, taking full advantage of the pair additivity of the potential. This model of PS was recently employed to calculate at the CG level the excess chemical potential of inserting additives in the polymeric matrix. The interaction between polymer and additives were also derived using CRW method. The agreement between results predicted employing CG and atomistic models is striking considering the sensibility of this thermodynamic property to the quality of the model.[71] The model has further been employed in hierarchical simulations of polystyrene surfaces.[72] This application of the derived PS model clearly requires a set of transferable potentials not biased to reproduce bulk-like chain packing characteristics (as would be the case for parameterized models).
In a recently published review by Karimi-Varzaneh *et al.*[73] the authors compare the performance of the different PS CG models, showing that the derived coarse graining method provides the only model able to reproduce the experimental density (at 1 atm.), the radial distribution function, the thermal expansion coefficient and the glass transition temperature of PS. The representability of this model is owing to the physical nature of the nonbonded pair potentials and the finer mapping scheme.

2.5.4 Ionic solutions

The ability to describe interactions between solvated ions is the first step to simulate biologically relevant systems. Obviously water is the main component in this

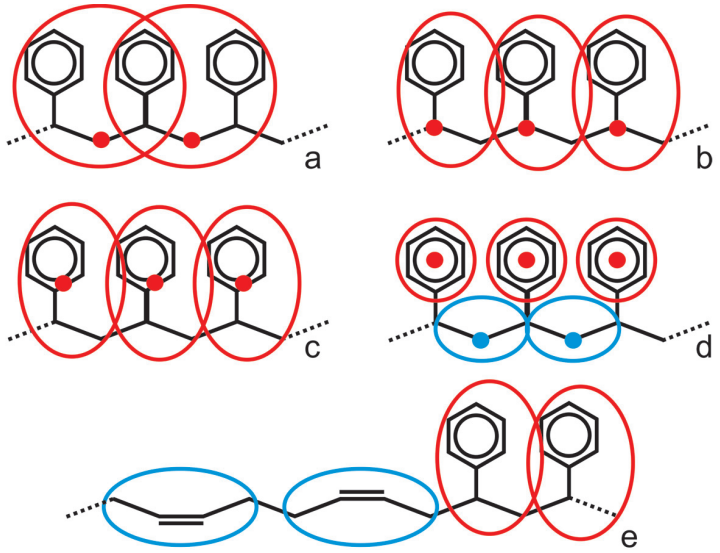


Figure 2.6: Mapping scheme of the polymers considered in section 2.5.3. The CG methods employed include: (a) IBI[63]; (b) IBI[65]; (c) IBI[41]; (d) CRW[18]; (e) IBI[70].

kind of systems and the opportunity of including the effect of water on the ion pairing by employing an implicit solvent model that captures the essential behavior of the system will enormously speed up the simulation while maintaining an accurate description of the system. In order to do this the CG pair potential between ions needs to be an *effective* pair potential. An appealing characteristic of a CG potential for ion will be its ability to work at different salt concentrations. Unfortunately to obtain this through an implicit solvent model is unlikely since the concentration of ions influences the water structure and this effect can not be easily plugged into a pair potential. Nevertheless it is possible in principle to obtain a series of potential to use at different concentrations, and to use them according to the system that is being studied.

Among the **parametrized** methods IMC has been extensively used to investigate the behavior of ionic solutions. Lyubatsev and Laaksonen developed *effective* potentials for implicit solvent sodium chloride solutions.[3] The authors investigated the differences between the interaction potentials obtained at different concentrations and PMFs calculated at the same concentrations. They found out that there is a small difference between the two which can be explained by the fact that at long distances the PMF decays as a screened Coulomb potential, while the interaction potential behaves as the primitive model [cite]. The authors also proved that the difference between the interaction potentials at different concentration is related to the variation of the dielectric constant of water at different salt concentration. Mirzoev and

Lyubartsev used this idea in a latter paper in order to estimate the variation of this property with the temperature finding a good agreement with the experimental and all atom simulation values.[74] IMC method has been successfully employed also to study with CG implicit solvent model the interaction between DNA and ions.[75]
Based on similar observation Hess *et al.*[20] developed a CG **derived** implicit solvent model for ions. Since multi body contributions has to be taken into account the potential is based on PMF calculations. The authors showed that at infinite dilution the short range PMF between the sodium and chloride ions is the sum of an *effective* interaction plus a Coulomb term, and that at long distances the interaction is purely Coulombic. The authors also presented a PMF calculated at a finite ion concentration, and proved that this one was simply a "shifted" version of the PMF at infinite dilution, and that the shifting was related to the variation of the coulombic interaction relative to the electrostriction effect that an increasing concentration of ions has on the water dielectric permittivity. It is therefore possible to use the interaction potential developed at infinite dilution in a wide range of concentrations by simply accounting for the variation of the dielectric constant. The authors confirmed this by calculating the ions excess coordination number and the osmotic coefficient up to a concentration up to 4.5M. At this high ions concentration the behavior of CG and reference atomistic simulations starts to deviate due to other multi-body effects that the model is no longer able to capture. Shen *et al.*[21] further characterized this model, showing that the quality of the interaction potential is crucial to properly describe such a sensitive property like the osmotic coefficient. They also characterized the pair structure of the ions in solution showing that the CG model that takes into account the variation of the dielectric permittivity is able to reproduce reasonably this property up to a concentration of few mol per liter. Li *et al.*[40] used the same approach to develop a solvent-free CG model for sodium polystyrene sulfonate (NaPSS) solutions. In their model, potentials of mean force between small molecules and ions, such as propane, benzene and methyl sulfonate ions were calculated to represent the effective CG nonbonded potentials between the corresponding CG beads. Their CG model successfully reproduces conformation details of the atomistic PSS chain, such as the distributions of CG bonded degrees of freedom, global stiffness of PSS chain with different tacticities quantified by the mean square distance between two CG beads separated from n bonds away, as well as the counter ion distributions around the polyelectrolyte chain. The obtained CG model was extended to systems with

different concentration of salt (NaCl). Furthermore, they also showed that this CG model can be easily transferred to partially sulfonated polystyrene aqueous solutions.

2.6 Concluding remarks

The development of CG methods is still an extremely active field and as have we shown different methods perform better when applied to different systems. Therefore there is not a unique answer that defines which method is the more convenient for CG a system. First it is first necessary to understand which interactions are relevant to the physics of the problem we want to investigate. Then it is possible to define a mapping scheme that takes care of these and at the same time it is possible to decide if the interaction potentials should be a "clean" pair interaction of an *effective* pair interactions. Sometimes this choice is enforced, but sometimes a wise choice can strongly enhance the representability and the transferability of the CG model. Finally the model has to be tested to check if it is effectively able to reproduce the desired properties and eventually if it is able to reproduce other properties. We should remind that a perfect agreement between thermodynamic response functions calculated at the atomistic level and at the CG level is almost impossible to obtain using effective CG methods due the simplified nature of any CG model. In this review we showed how parametrized CG model are able by construction to reproduce certain properties. Therefore in the case that are exactly known which properties of the atomistic systems are desired to be reproduced by the CG one, this class of methods delivers a smooth way to obtain such models. On the other hand we reported few examples based on the newer class of derived methods that proved the good representability and transferability of the CG model obtained with them. This was achieved preserving the characteristic interactions of the systems also in the CG model and this is possible only when the interaction potential have a clear physical meaning. Also this makes possible to understand in which conditions a model starts to fail in reproducing thermodynamic properties. Considering this systematically developed CG models looks promising in order to develop CG models.

2.7 Acknowledgment

This research was supported by the German Research Foundation (DFG) within the Cluster of Excellence 259 "Smart Interfaces – Understanding and Designing Fluid Boundaries"

2.8 Bibliography

[1] D. Fritz, K. Koschke, V. A. Harmandaris, N. F. A. van der Vegt, and K. Kremer. *Phys. Chem. Chem. Phys.*, 13:10412, 2011.

[2] F. Mueller-Plathe. *Chem. Phys. Chem.*, 3:754, 2002.

[3] A. Lyubartsev and A. Laaksonen. *Phys. Rev. E*, 52:3730, 1995.

[4] M. S. Shell. *J. Chem. Phys.*, 129:144108, 2008.

[5] F. Ercolessi and J. B. Adams. *Europhys. Lett.*, 26:583, 1994.

[6] S. Izvekov and G. A. Voth. *J. Phys. Chem. B*, 109:2469, 2005.

[7] J. W. Mullinax and W. G. Noid. *J. Phys. Chem. C*, 114:5661, 2010.

[8] B. Hess, C. Holm, and N. F. A. van der Vegt. *J. Chem. Phys.*, 124:164509, 2006.

[9] Y. Wang, W. G. Noid, P. Liu, and G. A. Voth. *Phys. Chem. Chem. Phys.*, 11:2002, 2009.

[10] E. Brini, V. Marcon, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 13:10468, 2011.

[11] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *J. Phys. Chem. B*, 111:7812, 2007.

[12] K. A. Maerzke and J. I. Siepmann. *J. Phys. Chem. B*, 115:3452, 2011.

[13] J/ R. Allison, S. Riniker, and W. F. van Gunsteren. *J. Chem. Phys.*, 136:054505, 2012.

[14] W. L. Jorgensen and J. Tirado-Rives. *J. Am. Chem. soc.*, 110:1657, 1988.

[15] M. G. Martin and J. I. Siepmann. *J. Phys. Chem. B*, 102:2569, 1998.

[16] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren. *J. Comput. Chem.*, 25:1656, 2004.

[17] W. Tschoep, K. Kremer, J. Batoulis, T. Burger, and O. Hahn. *Acta Polym.*, 49:61, 1998.

[18] D. Fritz, V. A. Harmandaris, K. Kremer, and N. F. A. van der Vegt. *Macromolecules*, 42:7579, 2009.

[19] D. Reith, H. Meyer, and F. Mueller-Plathe. *Comput. Phys. Commun.*, 148:299, 2002.

[20] B. Hess, C. Holm, and N. F. A. van der Vegt. *Phys. Rev. Lett.*, 96:147801, 2006.

[21] J. Shen, C. Li, N. F. A. van der Vegt, and C. Peter. *J. Chem. Theory Comput.*, 7:1916, 2011.

[22] P. Ganguly, D. Mukherji, C. Junghans, and N. F. A. van der Vegt. *J. Chem. Theory Comput.*, 8:1802, 2012.

[23] R. Henderson. *Phys. Lett. A*, 49:197, 1974.

[24] D. Reith, M. Puetz, and F. Mueller-Plathe. *J. Comput. Chem.*, 24:1624, 2003.

[25] W. Schommers. *Phys. Lett. A*, 43:157, 1973.

[26] W. Schommers. *Phys. Rev. A*, 28:3599, 1983.

[27] A. K. Soper. *Chem. Phys.*, 202:295, 1996.

[28] Z. Wang and M. Deserno. *J. Phys. Chem. B*, 114:11207, 2010.

[29] V. Ruehle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko. *J. Chem. Theory Comput.*, 5:3211, 2009.

[30] J. Sandeep, S. Garde, and S. K. Kumar. *Ind. Eng. Chem. Res.*, 45:5614, 2004.

[31] S. Jain, S. Garde, and S. K. Kumar. *Ind. Eng. Chem. Res.*, 45:5614, 2006.

[32] T. Murtola, E. Falck, M. Karttunen, and I. Vattulainen. *J. Chem. Phys.*, 126:075101, 2007.

[33] T. Murtola, M. Karttunen, and I. Vattulainen. *J. Chem. Phys.*, 131:055101, 2009.

[34] A. Chaimovich and M. S. Shell. *J. Chem. Phys.*, 134:094112, 2011.

[35] S. Izvekov and G. A. Voth. *J. Chem. Phys.*, 123:134105, 2005.

[36] W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen. *J. Chem. Phys.*, 128:244114, 2008.

[37] W. G. Noid, P. Liu, Y. Wang, J.-W. Chu, G. S. Ayton, S. Izvekov, H. C. Anderson, and G. A. Voth. *J. Chem. Phys.*, 128:244115, 2008.

[38] J. F. Rudzinski and W. G. Noid. *J. Chem. Phys.*, 135:214101, 2011.

[39] A. Villa, C. Peter, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 11:2077, 2009.

[40] C. Li, J. Shen, C. Peter, and N. F. A. van der Vegt. *Macromol.*, 45:2551, 2012.

[41] H. J. Qian, P. Carbone, X. Chen, H. A. Karimi-Varzaneh, C. C. Liew, and F. Mueller-Plathe. *Macromolecules*, 41:9919, 2008.

[42] C. R. Ellis, J. F. Rudzinski, and W. G. Noid. *Macromol. Theory. Simul.*, 20:478, 2011.

[43] K. Farah, A. C. Fogarty, M. C. Bohm, and F. Mueller-Plathe. *Phys. Chem. Chem. Phys.*, 13:2894, 2011.

[44] A. J. Rzepiela, M. Louhivuori, C. Peter, and S. J. Marrink. *Phys. Chem. Chem. Phys.*, 13:10437, 2011.

[45] V. Ruele and C. Junghans. *Macromol. Theory Simul.*, 20:472, 2011.

[46] F. Mueller-Plathe. *J. Chem. Phys.*, 106:6082, 1997.

[47] A. Lyubartsev, A. Mirzoev, L. J. Chen, and A. Laaksonen. *Faraday Discuss.*, 144:43, 2009.

[48] H. Wang, C. Junghans, and K. Kremer. *Eur. Phys. J. E*, 28:221, 2009.

[49] S. Izvekov and G. A. Voth. *J. Chem. Theory Comput.*, 2:637, 2006.

[50] W. G. Noid, J. W. Chu, G. S. Ayton, and G. A. Voth. *J. Phys. Chem. B*, 111:4116, 2007.

[51] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. *Intermolecular Forces*, pages 331–342. Reidel, 1981.

[52] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. *J. Chem. Phys.*, 79:926, 1983.

[53] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. *J. Phys. Chem.*, 91:6269, 1987.

[54] K. R. Hadley and C. McCabe. *J. Phys. Chem. B*, 114:4590, 2010.

[55] D. Steinley. *Br. J. Math. Stat. Psychol.*, 59:1, 2006.

[56] J. B. MacQueen. *Some Methods for Classification and Analysis of MultiVariate ObserVations, in 5th Berkeley Symposium on Mathematical Statistics and Probability*. 1967.

[57] S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth. *J. Chem. Phys.*, 120:10896, 2004.

[58] M. W. Mahoney and W. L. Jorgensen. *J. Chem. Phys.*, 112:8910, 2000.

[59] L. Larini, L. Lu, and G. A. Voth. *J. Chem. Phys.*, 132:164107, 2010.

[60] M. U. Hammer, T. H. Anderson, A. Chaimovich, M. S. Shell, and J. Israelachvili. *Faraday Discuss.*, 146:299, 2010.

[61] F. Mueller-Plathe. *Macromolecules*, 29(13):4782, 1996.

[62] W. L. Jorgensen and D. L. Severance. *J. Am. Chem. Soc.*, 112:4768, 1990.

[63] G. Milano and F. Mueller-Plathe. *J. Phys. Chem. B*, 109:18609, 2005.

[64] G. Milano, S. Goudeau, and F. Mueller-Plathe. *J. Polym. Sci. Part B Polym. Phys.*, 43:871, 2005.

[65] Q. Sun and R. Faller. *Computers and Chemical Engineering*, 29:2380, 2005.

[66] Q. Sun and R. Faller. *Macromolecules*, 39:812, 2006.

[67] T. Spyriouni, C. Tzoumanekas, D. Theodorou, F. Mueller-Plathe, and G. Milano. *Macromolecules*, 40(10):3876, 2007.

[68] C. Tzoumanekas and D. N. Theodorou. *Macromolecules*, 39:4592, 2006.

[69] L. J. Fetters, D. J. Lohse, D. Richter, T. A. Witten, and A. Zirkel. *Macromolecules*, 27:4639, 1994.

[70] X. Li, D. Kou, S. Rao, and H. Liang. *J. Chem. Phys.*, 124:204909, 2006.

[71] E. Brini, C. R. Herbers, G. Deichmann, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 14:11896, 2012.

[72] V. Marcon, D. Fritz, and N. F. A. van der Vegt. *Soft Matter*, 8:5585, 2012.

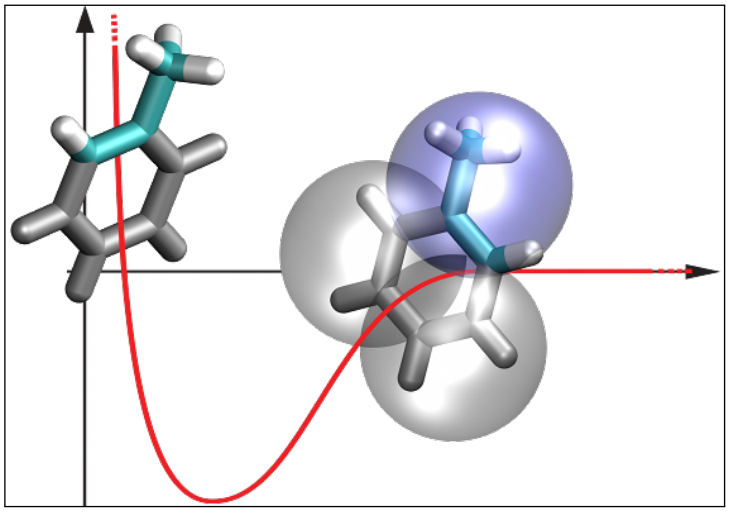
[73] H. A. Karimi-Varzaneh, N. F. A. van der Vegt, F. Mueller-Plathe, and P. Carbone. *Chem. Phys. Chem.*, page DOI: 10.1002/cphc.201200111, 2012.

[74] A. Mirzoev and A. P. Lyubartsev. *Phys. Chem. Chem. Phys.*, 13:5722, 2011.

[75] A. P. Lyubartsev and A. Laaksonen. *J. Phys. Chem*, 111(24):11207, 1999.

3 Conditional reversible work method for molecular coarse graining applications

Systematically coarse grained models for complex fluids usually lack chemical and thermodynamic transferability. Efforts to improve transferability require the development of effective potentials with unequivocal physical significance. In this paper, we introduce conditional reversible work (CRW) potentials that describe nonbonded interactions in coarse grained models at the pair level. The method used to obtain these potentials is straightforward to implement, can be readily extended to compute hydration contributions in implicit-solvent potentials, and is easy to automize. As a first illustration of the method, we present CRW potentials for 3-site models of hexane and toluene. The temperature-transferability of the liquid phase density obtained with these potentials has been investigated, and a comparison has been made with effective potentials obtained by the iterative Boltzmann inversion method.



3.1 Introduction

Systematically coarse grained models of complex fluids can be studied with computer simulations on time and length scales that reach well beyond scales accessible with detailed-atomistic models. In the past, several methodologies have been proposed to generating coarse grained (CG) models, a number of them requiring input from detailed atomistic simulations of the system the CG model intends to describe. Effective interactions between coarse grained sites can be obtained by bringing the pair interaction in consistence with the pair density[1–3] using automated methods such as the inverse Monte Carlo (IMC)[1] and Iterative Boltzmann inversion (IBI)[3] methods. Alternatively, thermodynamic data can be used to parameterize CG nonbonded potentials (whose functional form is arbitrarily chosen to be, e.g., of Lennard-Jones type) for sets of small molecules in the condensed phase,[4–7] following a philosophy also adopted in parameterization of all-atom (AA) force field models. Although the various methods have successfully been used to study biological and other soft condensed matter systems, several important challenges remain to be resolved and relate to questions concerning the coarse grained potentials. A first question addresses the dynamics of CG models which (in the absence of dissipative forces) is faster than AA models as a result of the potential energy surface being coarse grained. In particular in systems with multiple components, it is not a priori clear if time scales can be uniquely defined (see for example ref.[8]). A second question, which is being explored in this paper, relates to the chemical and thermodynamic transferability of the CG potentials. While AA force field models are usually state point dependent, limiting their validity to a small region of the phase diagram, the state dependence of CG models is even stronger. Because one CG state, specified by a set of CG degrees of freedom, corresponds to many realizations of the AA system, Boltzmann-weighted averaging over the AA realizations yields temperature and density-dependent potentials for the CG degrees of freedom. In practise, this state dependence is not well understood and will be different for potentials obtained through different methods.

The beads in the CG models of interest in this work typically merge 5-10 real atoms. Hence we may also refer to these beads as “superatoms” or “united atoms”. For a given system, the nonbonded potentials may be quite different depending on the coarse graining method that has been used to derive them and the target proper-

ties that have been selected to be reproduced with the model.[9] The functional form of a nonbonded pair potential that has been chosen ad hoc (e.g. a Lennard-Jones type) may be quite unphysical, despite that it has been successfully parameterized to reproduce several selected thermodynamic properties. On the contrary, no assumptions are made about the functional form of the pair potentials obtained by the IBI and IMC methods, but these potentials may nevertheless be unphysical, too. Despite the existence of a uniqueness relationship between the pair potential and the pair correlation function,[10] IMC- and IBI-derived pair potentials often show a quite complicated long range behavior (sometimes with several minima).[9] It could in fact be shown that several, quite different, effective pair potentials may reproduce the pair correlation functions within line thickness,[3] which probably reflects the fact that the liquid structure is largely determined by the short-range repulsive part of the potential, rather than by the long-range attractive part. An advantage emerging from this observation is that it provides additional degrees of freedom to "tune" other properties of interest, while keeping consistent structural properties.

In developing the CG models presented in this work, we follow a philosophy which avoids tuning the pair potential on target properties. To this end, we use reversible work calculations to map explicit atom potentials onto coarse grained potentials[11,12] and study the representability and temperature transferability of the resulting pair potential. The correct functional form, or distance-dependence, of the CG potential naturally results from reversible work approaches for any type of CG mapping, which is not necessarily true for IMC- and IBI-derived potentials. We here emphasize the functional form of CG potentials because it determines, amongst other properties, the statistical correlation between volume and enthalpy fluctuations in an NpT system, which, in turn, determines the thermal expansion coefficient of the fluid system[13] and therefore the temperature transferability of the model. This argumentation can be understood intuitively. Fluctuations in intermolecular distances cause volume and enthalpy fluctuations, with the enthalpy fluctuation being determined by the dependence of the pair potential on distance. In this paper we will show that these thermodynamic fluctuations are reproduced with CRW potentials in very good agreement with the atomistic model. As a result, the CRW models are temperature transferable in a relatively wide temperature range of up to 100 K. Temperature transferability of effective pair potential models, whose functional forms are chosen arbitrarily or are obtained as non-unique solutions by fitting against target proper-

ties, can not be warranted unless explicit parameterization on thermal expansion has entered the parameterization.[14]

In this paper, we propose a new method for deriving nonbonded pair potentials for CG models. The method is based on calculations of conditional reversible work (CRW) and differs with respect to a previous application of reverse work[11] as it does not perform iterations on the potentials. The CRW method does not parameterize against condensed phase properties obtained from atomistic simulations, hence properties like the liquid phase density and pair correlations functions naturally derive from the potential. The first application of the CRW method was reported by Fritz et al.[15] who developed a CG model for poly(styrene). In that work it was shown that the CG model could reproduce the liquid phase density of the polymer over a fairly broad range of temperatures in satisfactorily agreement with the atomistic model. Here, we apply the CRW method to hexane and toluene and examine the temperature transferability of the resulting CG potentials in molecular dynamics (MD) simulations of the liquid phase. A comparison is being made with CG potentials obtained with iterative Boltzmann inversion.

3.2 Conditional Reversible Work - CRW

In molecular coarse graining, atoms are merged into CG interaction sites, which interact through bonded potentials (if they are chemically linked within a molecule) and nonbonded potentials. The nonbonded potentials are assumed to be pairwise additive. To illustrate the concept of reversible work, Fig.3.1 shows a thermodynamic cycle in which the six connected spheres represent hexane in a united atom description. We are interested to map the two central methylene groups into a single CG interaction site and seek for an effective potential $U_{eff}(r)$ that describes the pair interaction between two of such CG sites at distance r. We *define* this potential as the reversible work expended upon introducing nonbonded interactions between the atoms that compose the two CG sites at distance r. Thus, the nonbonded interaction between the two hexane molecules in the upper right state in Fig.3.1 is unperturbed, while in the lower right state the nonbonded interactions between the central methylene groups of the two hexane molecules are “turned off” (which is here illustrated by representing the central methylene groups in white). $U_{eff}(r)$ can be obtained by calculating the free-energy change along the path indicated by the vertical arrow in

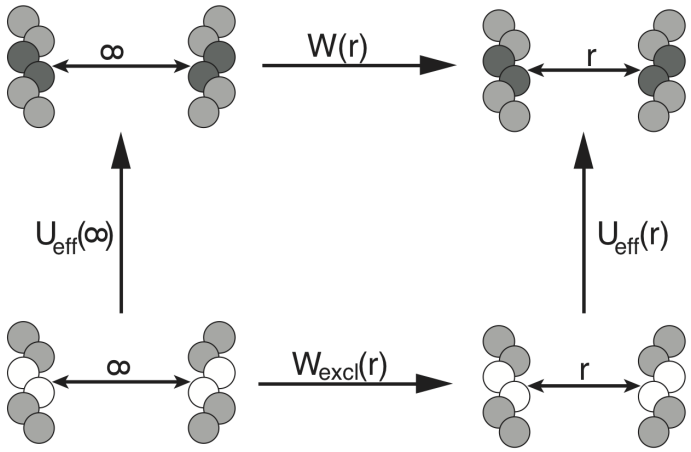


Figure 3.1: Thermodynamic cycle used to compute the conditional reversible work potential U_{eff}. In this example, the two central methylene groups in hexane are mapped on a single interaction site and the effective pair potential between two of these sites is calculated. To this end, the potential of mean force $W(r)$ is calculated with the atomistic force field in the upper part, where r denotes the distance between the centers of mass of the atoms composing the two sites. In the lower part, intermolecular interactions are excluded between the atoms merged in the central sites on the two molecules and the potential of mean force $W_{excl}(r)$ is calculated once more. The CRW potential is obtained from: $U_{eff}(r) - U_{eff}(\infty) = W(r) - W_{excl}(r)$.

Fig.3.1. However, by invoking the thermodynamic cycle shown in Fig.3.1, another scheme emerges, where the potential of mean force along r is calculated twice, i.e. once for the system with all interactions included ($W(r)$) and once for the system excluding direct interactions between the central methylene groups ($W_{excl}(r)$). We can therefore write

$$U_{eff}(r) = W(r) - W_{excl}(r) + U_{eff}(\infty) \tag{3.1}$$

$W(r)$ and $W_{excl}(r)$ can be calculated in an MD simulation, e.g., by sampling the mean force associated with the constraint r.

By means of this procedure, we compute reversible work $U_{eff}(r)$ to introduce interactions between two atom groups (two CG sites, one on each hexane molecule) in the natural chemical environment of the respective atom groups. We therefore impose a condition on the mutual orientations that can be adopted by these

two groups, which is determined by steric constraints of the surrounding chemical moieties. Hence, we refer to $U_{eff}(r)$ as *conditional* reversible work. To obtain CG nonbonded potentials for a 3-site hexane model, the above procedure is repeated, defining r between the two outer CH_2CH_3 groups of the two molecules and, once more, defining r between the outer CH_2CH_3 group and inner CH_2CH_2 group. If the CRW method is applied by considering only these two molecules in the calculation, furtheron denoted CRW(vacuum), the computational cost will be very small. Transferability of the resulting CRW(vacuum) potentials to the liquid phase at the same temperature will be explored below. CRW calculations may also be performed in the presence of a surrounding liquid medium, the resulting potentials are then denoted CRW(liquid). In this case, the procedure summarized in Fig.3.1 is performed in an atomistic environment of liquid hexane, i.e. $W_{excl}(r)$ is calculated with constraint MD in an atomistic hexane environment, while $W(R)$ is obtained from the pair correlation function, $g(r)$, of liquid hexane according to $W(r) = -k_B T ln[g(r)]$, where k_B is the Boltzmann constant and T the temperature. The CRW(vacuum) and CRW(liquid) potentials cannot assumed to be identical and differences between them are caused by different sampling spaces of the mutual orientation of the two solute molecules (Fig.3.1) in their respective (nonbonded) environments. Irrespective of this, the CRW potentials keep their significance of reversible work expended upon introducing nonbonded interactions between the atoms that compose two CG sites at distance r. CRW(vacuum) and CRW(liquid) potentials shall both be examined in this paper for hexane. A recent CRW(vacuum) potential for poly(styrene) melts[15] showed good transferability properties, providing an accurate description of the melt density in NpT simulations for various temperatures. We surmise that this successful application of CRW(vacuum) potentials in the liquid phase is owing to the inclusion of the chemically bonded environment in the sampling of the CRW(vacuum) potentials, since the chemical environment immediately adjacent to the CG beads (which is identical in vacuum and in the liquid) has the largest effect on the CG pair interactions.

3.2.1 Implicit solvent CRW potentials

The CRW method can easily be generalized to obtain implicit-solvent potentials for coarse grained solutes in aqueous solution. In this case, the potential of mean force calculation in the lower part of Fig.3.1 can be performed with additional exclusion of

the nonbonded interactions between the central methylene groups and the solvent molecules, in which case the thermodynamic cycle provides an implicit-solvent potential $U_{eff}(r)$. Note that in this case $U_{eff}(r)$ includes the CRW due to the "direct" nonbonded interactions between the atoms that compose the two CG sites *and* the *conditional hydration free energy* of the two sites (relative to the conditional hydration free energy with the two sites at large separation). Conditional hydration refers to the fact that the atoms being hydrated are in their natural chemical environment of the solute molecule. We will further explore the CRW method for deriving implicit solvent potentials in a forthcoming paper.

3.3 Computational details

Toluene was modelled with the OPLS-AA force field,[16] while for hexane the OPLS-UA force field[17] was used. In all atomistic simulations, we used a cut-off distance of 1.0 nm for the electrostatic interactions, while for the Lennard-Jones interactions we used a twin-range cut-off scheme with 1.0 nm and 1.3 nm cut-off radii. Neighbourlist updates were made every 10 time steps. A long range dispersion correction was applied for the energy. If not otherwise specified, the weak coupling method was used to maintain temperature and pressure constant.[18] The temperature coupling time was 0.1 ps, while the pressure coupling time was 1.0 ps with an isothermal compressibility equal to $4.5 \cdot 10^{-5}$ bar^{-1}. Cubic simulation boxes with 1000 molecules were used in all simulations of liquid hexane and toluene. 2 ns trajectories were accumulated. All simulations have been carried out with the GROMACS program.[19,20]

CRW(vacuum) potentials were obtained by performing a series of n distance constraint simulations between the CG mapping points using the linear constrain solver (LINCS) algorithm.[21,22] The average constraint force of every simulation is then integrated over the constraint coordinate in order to get the reversible work.[23] The constraint distance was varied between 1.30 nm and 0.32 nm in steps of 0.02 nm. The constraint dynamics simulations were all performed at 300 K for 400 ns, using the stochastic dynamics integrator of GROMACS with a friction constant of 0.5 ps^{-1} and a time step of 2 fs. CRW(liquid) potentials were obtained using constraint distances between 1.10 nm and 0.32 nm in steps of 0.02 nm.

The IBI potentials were calculated using the VOTCA toolkit.[9] Target radial distribution functions (RDFs) between the CG mapping points were obtained from NpT AA-simulations of 2ns at 298 K and 1 atm. The CG simulations during the IBI procedure were carried out in the NVT ensemble using a stochastic dynamics integrator

at a temperature of 298 K with a friction coefficient of 0.2 ps^{-1}. These simulations are 100 ps long with a timestep of 1 fs. During the iterative procedure the first 65 iterations are done in order to converge the radial distribution function. In the next 15 iterations the potential is updated with the double target of the radial distribution function and the pressure of 1 atm.[9] The last 115 iterations are done in order to converge the pressure without updating anymore the correction due to the radial distribution function. Since the pressure oscillates around the target value the chosen set of interaction parameters is the one that in this last series reproduce better the target pressure of 1 atm.

Coarse grained MD simulations were carried in the NpT ensemble using the stochastic dynamic integrator with a time step of 2 ps, a friction coefficient of 0.2 ps^{-1}, and the Parrinello-Rahman barostat.[24,25] The cutoff distance was 1.3 nm for the IBI and CRW(vacuum) potentials and 1.1 nm for the CRW(liquid) potentials. Since all CG beads are electrically neutral, no Coulombic interactions are present in the CG systems. The effective potentials obtained using the CRW and IBI methods were used in tabulated form.

3.4 Results and discussion

3.4.1 Mapping schemes

For hexane we used a 3 bead mapping scheme, merging 2 united atoms in every bead (Fig.3.2, left). Hence, we obtain two different bead types: type A represents the CH_3-CH_2- group, and type B represents the $-CH_2-CH_2-$ group. The potentials for the $A-B$ bond and $A-B-A$ angle were obtained by Boltzmann inverting the angle and bond length distribution functions calculated from an atomistic simulation of one molecule in vacuum. Toluene is a rigid aromatic molecule with an anisotropic shape, which we tried to represent using a 3 bead mapping scheme (Fig.3.2, right). The two yellow beads merge the same set of atoms (2 neighboring C-atoms of the ring together with the connecting H-atoms), but their environment is different. In fact they are linked to a primary or to a secondary carbon of the atoms that compose the gray bead. Therefore we will treat these two CG beads as different species (B and C), although we do not expect a dramatic difference in their interaction potentials. The remaining bead (A) represents the toluene methyl group together with two aromatic carbons and one aromatic hydrogen. Although it is not evident from this mapping scheme that the coarse grained potential will have a C_2 axial symmetry, we note

that symmetry aspects are always accounted for in systematic coarse graining methods. The CRW method used here takes into account the direct chemical environment within the molecule in obtaining effective potentials for the CG beads. Through that approach, information on the molecular symmetry enters into the potentials. Considering the rigidity of the toluene molecule, the intramolecular distances between the coarse-grain beads are constrained at the equilibrium distances between the mapping points.

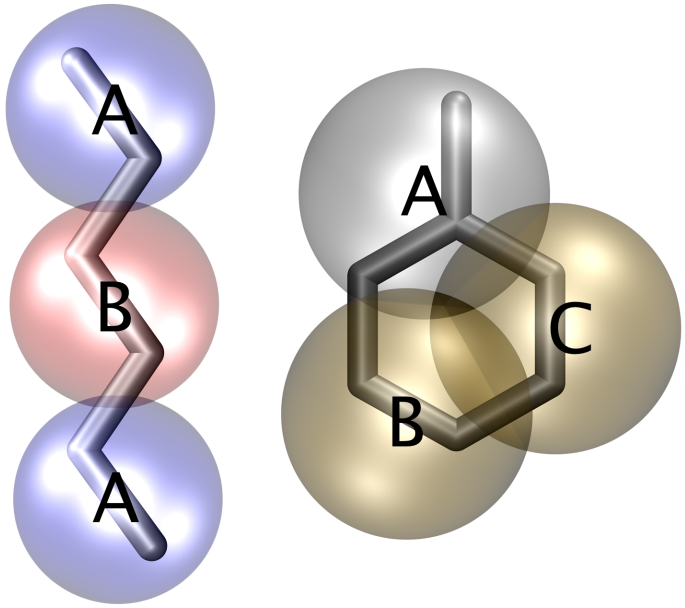


Figure 3.2: Mapping scheme for the CG representation of hexane (left) and toluene (right). In this work, the "atomistic model" used for hexane corresponds to a six-site united-atom model. The atomistic model used for toluene is 15-site all-atom model. The hydrogen atoms are not displayed here for clearness. The figure was generated using VMD. [26]

3.4.2 Hexane

Figure 3.3(a) shows $U_{eff}(r)$ for the A-A interaction in hexane together with $W(r)$ and $W_{excl}(r)$. The latter quantities show several inflection points for distances beyond the location of the minimum. These are due to 'indirect' interactions involving the atoms not belonging to the A-beads between which the distance constraint was set. By construction, these indirect contributions do not appear in the CRW potential, i.e. they are removed by applying Eq 1, hence $U_{eff}(r)$ contains only the contribu-

tions of the direct interactions. The effective A-A, B-B, and A-B potentials for hexane are shown in Fig.3.3(b). The three potentials behave as expected. The A-A potential has a slightly bigger excluded volume and a deeper minimum than the B-B potential. This reflects the bigger size and the stronger interaction due to the presence of the terminal CH_3 groups. The strength of the A-B cross interaction falls in between A-A and B-B.

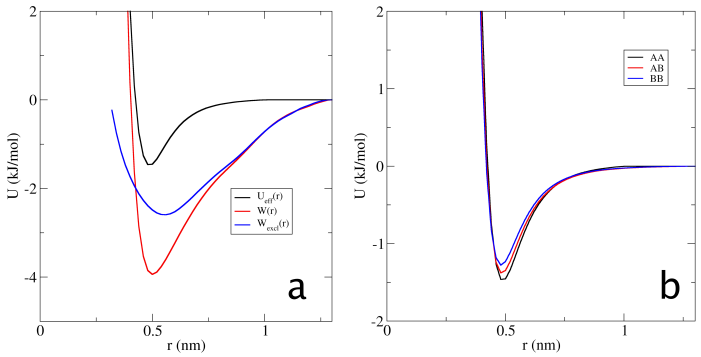


Figure 3.3: (a) The CRW(vacuum) potential U_{eff} for the A-A interaction in hexane together with the potentials of mean force $W(r)$ and $W_{excl}(r)$ (cf. Fig.3.1) at T=298 K. (b) CRW(vacuum) potentials for hexane, T=298 K.

The mass densities of bulk hexane obtained with the CRW models are shown in Fig.3.4 together with the mass density of the OPLS-UA model used for developing the CG potentials. The densities were obtained from NpT simulations (1 atm) using the CRW potentials developed at 298 K and are presented for 5 different temperatures in order to validate the temperature transferability of the CRW models against the OPLS-UA model. Mass densities obtained with the CRW(vacuum) model are systematically too high (around 9.5%), while the densities obtained with the CRW(liquid) model are in good agreement with the parent OPLS-UA model. The temperature dependence of the mass density shows very good agreement between the CRW and OPLS-UA models. Hence, the CRW potentials are temperature transferable. A recently developed IBI model of liquid hexane by Farah et al.[14] (which is based on the same CG mapping scheme) shows, in contrast to our result, a rather poor temperature transferability, which can be improved a posteriori with temperature dependent scaling factors applied to the potentials. We will come back to the temperature transferability later on in this paper. The thermal expansion coefficients of the atomistic and CRW models are summarized in table 3.1.

The liquid structure characterized by the radial distribution function (RDF) of the hexane center of mass is shown in Fig.3.5. Comparison of the RDFs corresponding to

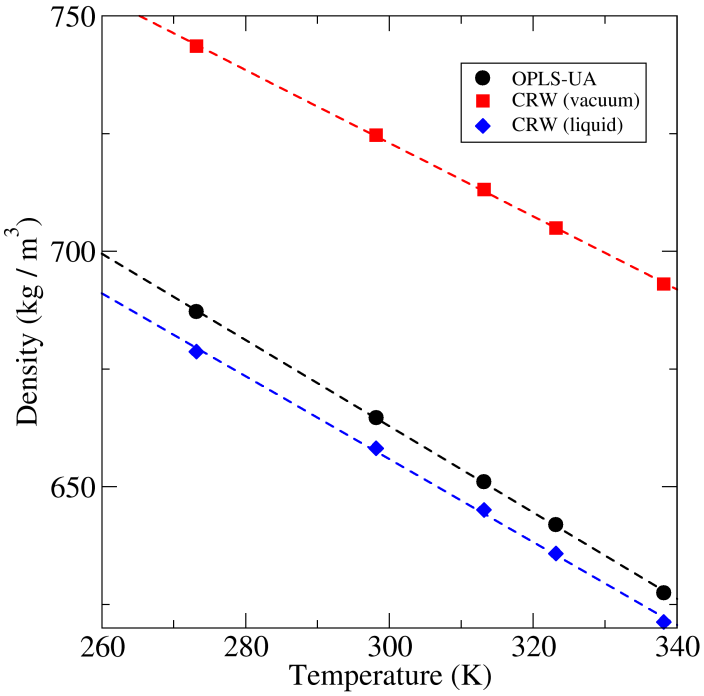


Figure 3.4: Mass density of liquid hexane as a function of temperature obtained from NpT simulations (1 atm.) with the atomistic (OPLS-UA) and CRW models.

Hexane	OPLS-UA	CRW (vacuum)	CRW (liquid)
α_p (K^{-1})	$1.39 \cdot 10^{-3}$	$1.08 \cdot 10^{-3}$	$1.35 \cdot 10^{-3}$
Toluene	OPLS-AA	CRW (shifted)	
α_p (K^{-1})	$1.40 \cdot 10^{-3}$	$1.27 \cdot 10^{-3}$	

Table 3.1: Thermal expansion coefficients for atomistic- and CRW models of hexane (vacuum and liquid sampling) and toluene.

the CRW and OPLS-UA models shows discrepancies in the peak heights. Apart from these differences, the overall agreement is satisfactory.

3.4.3 Toluene

Because of its rigid anisotropic shape toluene is a more challenging molecule to coarse-grain compared to hexane. Although shape anisotropy may be retained with a 3-bead representation, we still map sp2-carbons (and their surrounding atoms) on a spherically symmetric potential. Depending on the relative orientation of the two molecules, the excluded volume repulsion will be different for a given distance r between the CG sites. Fig.3.6 A and B show a side-to-side and shifted-stack configuration. Steric interactions between the hydrogen atoms lead to larger excluded

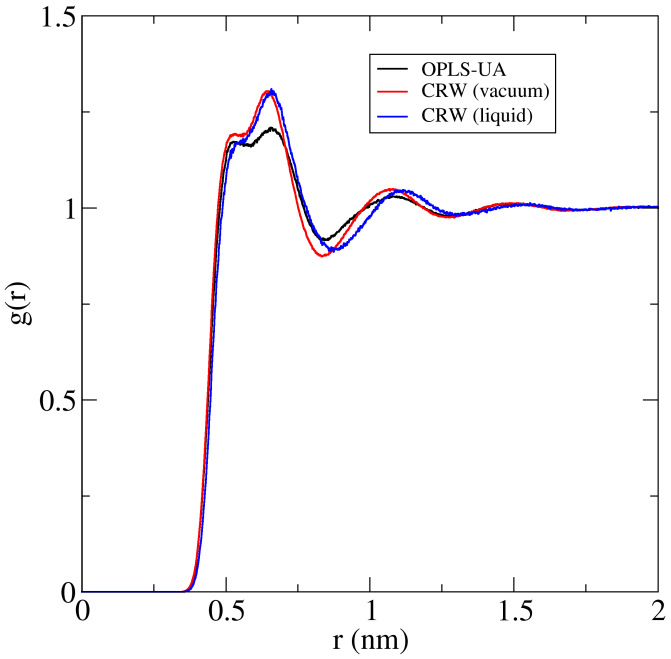


Figure 3.5: Molecular center of mass RDFs in liquid hexane at 298.15 K and 1 atm. OPLS-UA model (black) and CRW(vacuum) model (red) and CRW(liquid) model (blue).

volume repulsion in the side-to-side configuration as compared to the shifted-stack configuration. These configurations contribute to the CRW potential with the Boltzmann weight of their energy. But independent of the relative weights, the distance where the repulsion sets in in the CRW potential will always be underestimated for side-to-side configurations. In Fig.3.6 C it can be seen that in the plane of the ring the excluded volume of the CG model is too small to describe toluene side-to-side interactions. We tried to recover the correct excluded volume by shifting the positions of the CG beads from the center of mass (COM) to the center of geometry (COG) of the groups of atoms they represent. Note that we keep the CRW potentials obtained with the beads located at their centers of mass. The CG model obtained by this shifting procedure resembles the atomistic molecule much better as illustrated in Fig.3.6 D. We emphasize that this shifting procedure is necessary, because independently from the mapping point chosen to calculate the potentials (COM or COG of the group of atoms) the problem shown in Fig.3.6 A and B is always present.

This is confirmed by considering the liquid density in NpT simulations using CRW(vacuum) potentials (Fig.3.7). The density obtained with the 2 non-shifted models is approximately 30% too high compared with the OPLS-AA model. The model that employs the potentials calculated in the center of geometry (CRW_{COG})

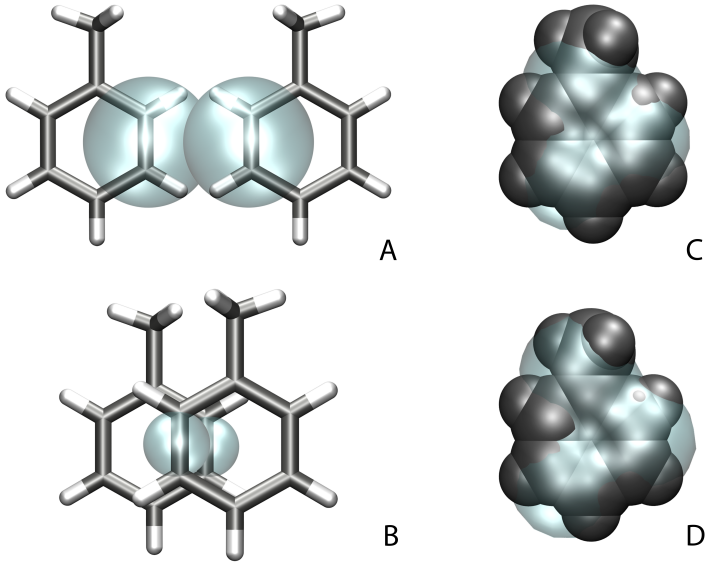


Figure 3.6: The excluded volume interaction of toluene depends on the orientation of mutual approach. (A) the side-to-side configuration, (B) the shifted-stack configuration. (C) quantitative representation of the excluded volume of the CRW model (light blue). The excluded volume is underestimated compared with the excluded volume of the atomistic model represented by the Van der Waals radii of the atoms composing the molecule (black). An improved representation of the excluded volume is obtained by shifting the positions of the CG beads from the COM to the COG of the atoms they represent (D).

however predicts a slightly lower density, indicating that the position of the interaction site is important in order to reproduce the behavior of anisotropic molecules. The model that has the positions of the beads shifted from the COM to the COG performs better than the non-shifted models. The predicted density of this model deviates on average only 4% from the OPLS-AA model (Fig.3.7). This model predicts also a thermal expansion coefficient much closer to the one of the OPLS-AA model (table 3.1).

3.4.4 Comparison between CRW and IBI models of toluene

In addition to CRW(vacuum) potentials, we also computed IBI potentials for toluene based on the same CG mapping scheme (Fig.3.2). The potentials are shown in Fig. 3.8. In contrast to the CRW(vacuum) potentials, the IBI potentials contain multiple minima. This can be due to a "non complete convergence" of the iteration process or be due to multibody contributions entering the pair potential. It is interesting to consider the BB and CC potentials. The CRW(vacuum) potentials for the

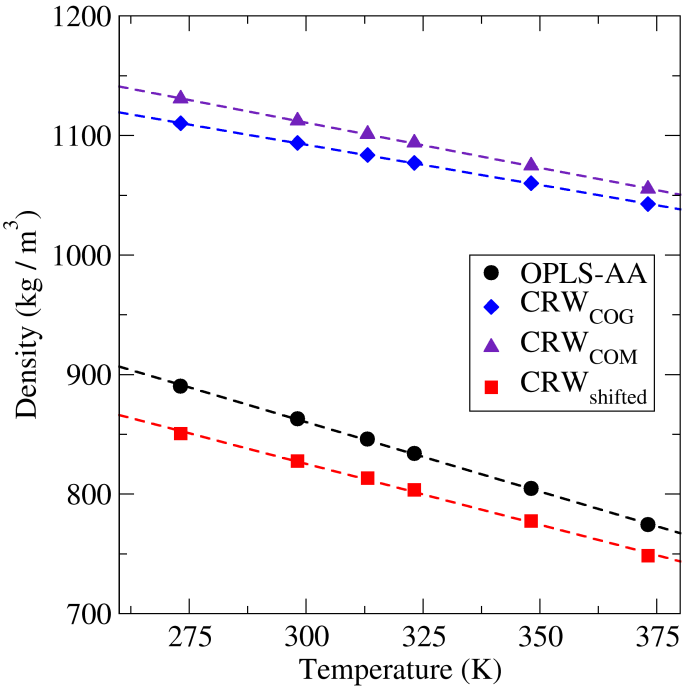


Figure 3.7: Mass density of liquid toluene as a function of temperature obtained from NpT simulations (1 atm.) with the atomistic (OPLS-AA) and CRW(vacuum) models (COM, COG and shifted).

BB and CC interactions in Fig. 3.8 (top panel, right) differ only marginally, which is expected because the B and C groups merge the same atoms. However, the chemical environment of these groups is different, which causes differences in the BB and CC RDFs of the liquid. The BB and CC potentials obtained with IBI are not identical (lower panel, right), indicating that the final pair potential includes a part of the indirect many body correlations. The same occurs for the AB, AC, and BC IBI potentials.

The IBI potentials were developed using the RDFs and the pressure ($1atm.$) as targets. When pressure is included in the iteration process, the long-range part of the potential converges faster. The long-range part usually has no strong effect on the RDF, therefore inclusion of a thermodynamic target quantity (for example pressure) accelerates the convergence of the IBI procedure. It should be noted, however, that a unique correspondence between the pressure and the potential does not exist, hence many potentials may in principle yield the same pressure. With a multi-bead mapping scheme, like the one employed here for the toluene, multiple potentials contribute to the pressure, and errors in one of the potentials may be cancelled by errors in another potential. The IBI potentials are perfectly able to reproduce the target properties (RDFs between the coarse grain mapping points and pressure),

but considering their long-range tail behaviour we cannot expect that they show a good thermodynamic transferability. To predictively model the liquid state density at another temperature (thermal expansion) it is important to correctly describe the correlation of volume and enthalpy fluctuations in the constant NpT system; i.e. $\alpha_p \sim \langle \delta V \delta H \rangle_{NpT}$ with $\delta X = X - \langle X \rangle_{NpT}$ denoting the instantaneous fluctuation of observable X. When volume fluctuates, the distances between the molecules in the system fluctuate as well. Accordingly, different regions of the intermolecular potential are sampled, which gives rise to the corresponding enthalpy fluctuation. Only when the dependence of U_{eff} on r is physically realistic, we expect a correct representation of $\langle \delta V \delta H \rangle_{NpT}$.

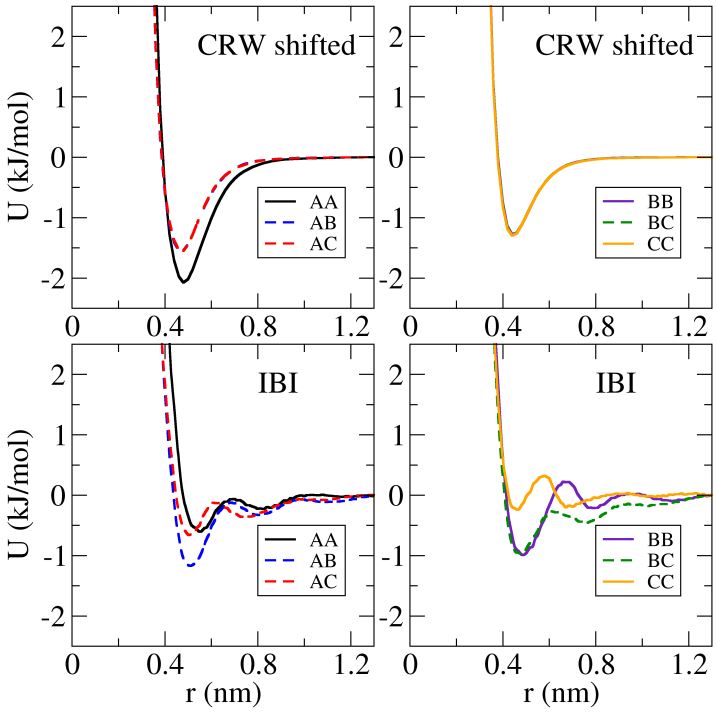


Figure 3.8: CRW(vacuum) potentials (top panel) and IBI potentials (bottom panel) for the 3-site toluene model in Fig.3.2.

Because we expect the shape of the potentials to relate to the temperature transferability of the models, we compared the liquid phase density obtained at different temperatures for the IBI and CRW toluene models. To this end, we investigated the shifted CRW model and two IBI models, one developed using the COMs as mapping points and another using the COGs of the atoms composing the CG sites as mapping points. The liquid phase density obtained with these models is shown in Fig. 3.9 together with the data obtained with the OPLS-AA model. The CRW and IBI potentials were developed at 298 K and used in simulations at lower and higher temperatures.

By construction, the IBI potentials reproduce the OPLS-AA density at 298 K. Clearly, however, the thermal expansion coefficients of the IBI models are significantly too low.

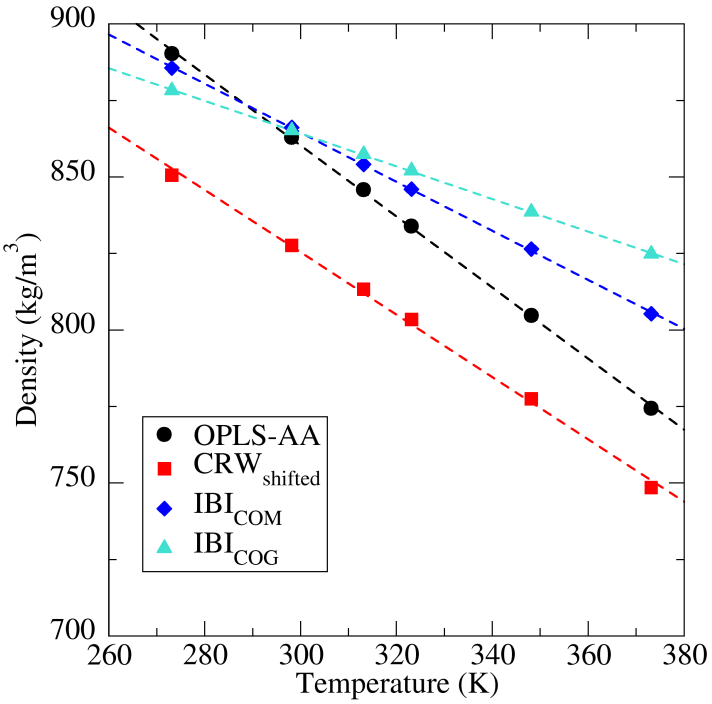


Figure 3.9: Mass density of liquid toluene as a function of temperature obtained from NpT simulations (1 atm.) with the atomistic model (OPLS-AA), shifted CRW(vacuum) model, and the IBI models developed using the COM and COG of the mapped atoms.

The RDFs of the molecular centers of mass obtained with the CG models are shown in Fig.3.10. For comparison the RDF of the OPLS-AA model is included as well. The RDF obtained with the CRW model shows reasonable agreement with the reference RDF (OPLS-AA model), however, the maxima and minima could not be reproduced.

3.5 Discussion and Conclusions

Systematically coarse grained potentials are necessarily dependent on thermodynamic conditions and environment. Potentials obtained by inverse Monte Carlo,[1] iterative Boltzmann inversion,[3,27] (conditional) reversible work, or by other methods[4–7,28,29] are state dependent, the extent of which has been investigated in this paper by comparing CRW and IBI potentials. We have shown that CG potentials obtained with the CRW method provide encouraging results for liquid phase properties of two molecular liquids. It could in particular be demonstrated that the tempera-

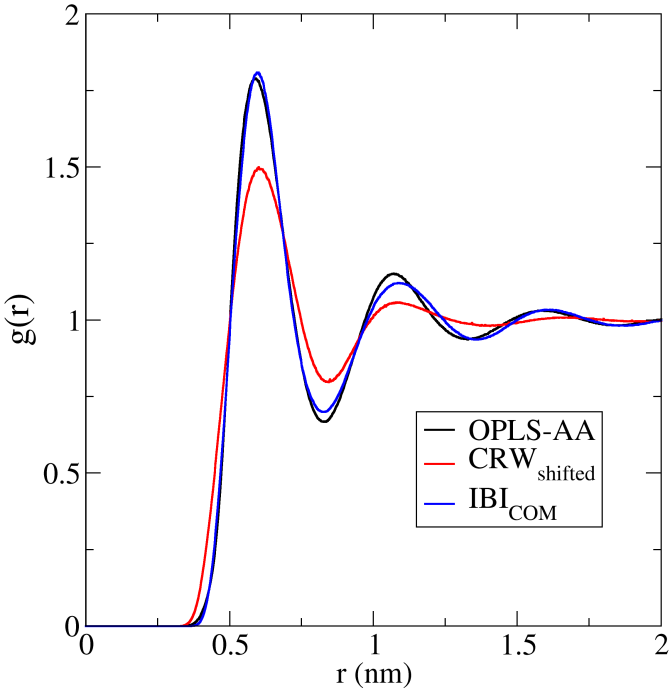


Figure 3.10: Center of mass RDFs of liquid toluene at 298.15 K and 1 atm. obtained with the OPLS-AA, CRW(vacuum), and IBI models. Note that the COM RDF has not been used as a target function in IBI.

ture transferability of the CRW-derived potentials is in reasonable agreement with the atomistic model. CRW potentials are distinctly different from potentials obtained by alternative coarse graining methods in the sense that they are physically inspired, i.e. the CRW potential is a free energy function, which unambiguously relates to the interatomic interactions between only those atoms merged within a pair of coarse grained beads. Hence, the long-range attractive tail of the CRW potential is monotonically varying with distance, unlike the tails of the potentials obtained by IBI or force matching, as was demonstrated previously.[9] We believe that this property of the CRW potential is important in future development of CG potentials with improved temperature transferability. It is interesting to note that upon completion of this work a recent study on hybrid simulations by Rzepiela et al.[30] was brought to our attention, in which the authors describe a single-site butane potential, developed based on ideas similar to the ones used here to develop the CRW(liquid) pair potentials for the 3-site hexane model. Rzepiela et al. showed that the 1-site butane potential closely reproduces the RDF of the liquid obtained with the atomistic model and is suited to be combined with the atomistic force field model, i.e. coarse grained and atomistic interactions can be "mixed" to yield liquid butane properties in reasonable agreement with the full atomistic model.

The question of chemical transferability (as opposed to thermodynamic transferability) has not been addressed in this work. This will be topic of a future work, but at this point a few comments are appropriate. A CRW potential for a pair of CG sites is obtained, satisfying the condition that these sites are embedded within the chemical environment of the overall molecule considered. This condition ensures that in the derivation of the potential, the two CG sites adopt only those mutual orientations that do not violate steric constraints imposed by their immediate chemical environment. A previous application to a macromolecular system, in which the nonbonded CRW potentials were obtained from smaller fragments, provided a temperature transferable model in good agreement with the corresponding atomistic model.[15] Because the CRW method is computationally cheap and easily automized, CRW potentials for specific atom groups in various chemical environments typical of a larger class of organic and biological molecules can in principle be obtained. We finally point out that the CRW method can also be used to derive implicit solvent potentials, e.g., for biological peptides or polyelectrolytes in water, accounting for the conditional hydration free energy in the effective pair potential.

3.6 Acknowledgements

We thank Christine Peter, Christoph Junghans and Victor Rühle for stimulating discussions and assistance with the VOTCA package.

3.7 Bibliography

[1] A. P. Lyubartsev and A. Laaksonen. *Phys. Rev. E*, 52:3730, 1995.

[2] R. L. C. Akkermans and W. J. Briels. *J. Chem. Phys.*, 114:1020, 2001.

[3] D. Reith, M. Putz, and F. Müller-Plathe. *J. Comput. Chem.*, 24:1624, 2003.

[4] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein. *J. Phys. Chem. B*, 105:4464, 2001.

[5] Siewert J Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H de Vries. *J Phys. Chem. B*, 111:7812, 2007.

[6] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein. *J. Chem. Phys.*, 119:7043, 2003.

[7] S. J. Marrink, A. H. de Vries, and A. E. Mark. *J. Phys. Chem. B*, 108:750, 2004.

[8] D. Fritz, C. R Herbers, K. Kremer, and N. F. A van der Vegt. *Soft Matter*, 5:4556, 2009.

[9] V. Ruehle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko. *J. Chem. Theory Comput.*, 5:3211, 2009.

[10] R. L. Henderson. *Phys. Lett. A*, 49:197, 1974.

[11] J. D. McCoy and J. G. Curro. *Macromolecules*, 31:9362, 1998.

[12] H. Fukunaga, T. Aoyagi, J. Takimoto, and M. Doi. *Comp. Phys. Comm.*, 142:224, 2001.

[13] A. Ben-Naim. *Molecular theory of solutions*. Oxford University Press, New York, 2006.

[14] K. Farah, A. C. Fogarty, M. C. Böhm, and F. Müller-Plathe. *Phys. Chem. Chem. Phys.*, 13:2894, 2011.

[15] D. Fritz, V. A. Harmandaris, K. Kremer, and N. F. A van der Vegt. *Macromolecules*, 42:7579, 2009.

[16] W. L. Jorgensen, D. S. Maxwell, and J. TiradoRives. *J. Am. Chem. Soc.*, 118:11225, 1996.

[17] W. L. Jorgensen and J. TiradoRives. *J. Am. Chem. Soc.*, 110:1657, 1988.

[18] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. *J. Chem. Phys.*, 81:3684, 1984.

[19] H. J. C. Berendsen, D. Vanderspoel, and R. Vandrunen. *Comput. Phys. Commun.*, 91:43, 1995.

[20] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Theory Comput.*, 4:435, 2008.

[21] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. *J. Comput. Chem.*, 18:1463, 1997.

[22] B. Hess. *J. Chem. Theory Comput.*, 4:116, 2008.

[23] A. Villa, C. Peter, and N. F. A. van der Vegt. *J. Chem. Theory Comput.*, 6:2434, 2010.

[24] M. Parrinello and A. Rahman. *J. Appl. Phys.*, 52:7182, 1981.

[25] S. Nose and M. L. Klein. *Mol. Phys.*, 50:1055, 1983.

[26] W. Humphrey, A. Dalke, and K. Schulten. *J. Mol. Graph.*, 14:33, 1996.

[27] P. Carbone, H. A. Karimi-Varzaneh, X. Chen, and F. Müller-Plathe. *J. Chem. Phys.*, 128:064904, 2008.

[28] S. Izvekov and G. A. Voth. *J. Chem. Phys.*, 123:134105, 2005.

[29] W. G. Noid, J. W. Chu, G. S. Ayton, and G. A. Voth. *J. Phys. Chem. B*, 111:4116, 2007.

[30] A. J. Rzepiela, M. Louhivuori, C. Peter, and S. J. Marrink. *Phys. Chem Chem. Phys.*, 13:10437, 2011.

3.8 Supplementary Informations

3.8.1 Symmetry of the CG interaction potential with respect to the original symmetry of the molecules

In section 3.4.1 We briefly acknowledged that in systematic CG even when the mapping schemes does not respect the molecular symmetry the sum of the interaction potential expresses it correctly. To prove that we calculated the energy for a CG system where a probe (CG toluene A bead) was moved perpendicular to the atomistic symmetry plane. In Fig. 3.11 are shown the results for this kind of analysis carried for different heights of the probe (Z_{probe}) with respect to the plane of the toluene molecule.

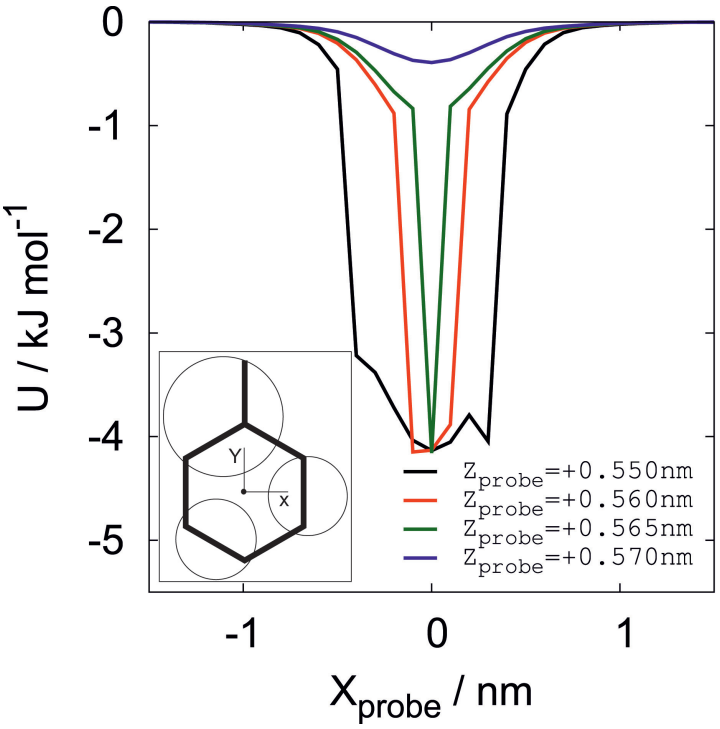
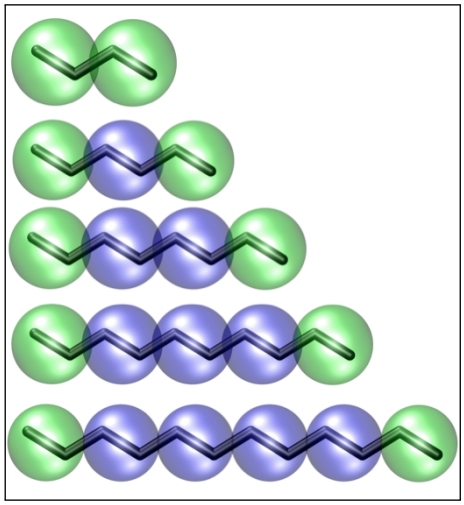


Figure 3.11: Scan of the potential energy of a CG system composed of a toluene molecule and a probe for different probe heights (Z_{probe}) from the plane of the phenyl ring. The probe is moved perpendicularly to the symmetry plane of the atomistic representation of toluene (X_{probe}). The insert reports the reference coordinate system; for all the scans $Y_{probe} = 0$. When the probe sits further away than a certain distance the sum of the CG interaction potentials reproduces the original symmetry of the atomistic molecule. At shorter distances this is not strictly true anymore, but the number of molecules that sample this region is small.

The symmetry of the interaction energy between the probe and the CG molecule brakes only when the two are at close distance. By looking at the RDFs of toluene (Fig. 3.10) it is possible to notice that this region of distances is sampled by only few molecules. Therefore we can now confirm that the symmetry of the original molecule is respected also by CG models with a mapping scheme that brakes the original symmetry.

4 Chemically transferable coarse-grained potentials from conditional reversible work calculations

The representability and transferability of effective pair potentials used in multiscale simulations of soft matter systems is ill understood. In this paper, we study liquid state systems composed of n-alkanes, the coarse-grained (CG) potential of which may be assumed pairwise additive and has been obtained using the conditional reversible work (CRW) method. The CRW method is a free-energy-based coarse-graining procedure, which, by means of performing the coarse graining at pair level, rigorously provides a pair potential that describes the interaction free energy between two mapped atom groups (beads) embedded in their respective chemical environments. The pairwise nature of the interactions combined with their dependence on the chemically bonded environment makes CRW potentials ideally suited in studies of chemical transferability. We report CRW potentials for hexane using a mapping scheme that merges two heavy atoms in one CG bead. It is shown that the model is chemically and thermodynamically transferable to alkanes of different chain lengths in the liquid phase at temperatures between the melting- and the boiling point under atmospheric (1 atm) pressure conditions. It is further shown that CRW-CG potentials may be readily obtained from a single simulation of the liquid state using the free energy perturbation method, thereby providing a fast and versatile molecular coarse graining method for aliphatic molecules.



4.1 Introduction

Pair potentials for classical coarse-grained (CG) models of condensed fluid phases have been obtained in recent years by means of different coarse graining approaches. Here we shall be interested in CG models in which atoms of selected chemical groups within a molecule are merged and treated as a single interaction site. While united atom models that merge hydrogen atoms with aliphatic carbon atoms to which they are covalently bound are well-known, we will here be interested in models in which this coarse graining approach is taken a step further via a procedure which we refer to as systematic coarse graining.

A systematic coarse graining procedure may be viewed to consist of two steps: (i) defining a mapping scheme that connects the fine-grained (*e.g.* atomistic) and coarse-grained systems and (ii) calculating the potential of mean force that describes the interaction between the CG degrees of freedom. The potential of mean force is obtained by sampling the configuration space of the fine-grained model and corresponds to the free energy of the eliminated atomistic degrees of freedom. Since this procedure yields a multibody potential, which is per se not very useful in practical applications, different systematic coarse graining methods for developing *effective* pair potentials have been reported in recent years[1–11].

Molecular dynamics simulations with systematically coarse-grained models considerably extend the accessible time and length scales in comparison to atomistic models due to a reduction in the amount of computer time needed every force calculation step, the possibility to use a larger integration time step and a reduction of the effective friction between the particles which leads to a faster exploration of phase space.[12] However, these CG models usually lack chemical and thermodynamic transferability. Chemical transferability is achieved when the interaction potential for a specific chemical group can be used to model different molecules containing that group, while thermodynamic transferability is achieved when the potential can be used at different thermodynamic state points (temperature, pressure, mixture composition). Chemical transferability can however only be obtained when the effective pair potentials are physical, *i.e.* when they realistically describe the distance dependent, net attractive and repulsive interactions between the chem-

ical moieties mapped into the CG interaction sites.

In this paper, we study the chemical and thermodynamic transferability of CG alkane models obtained by applying a recently introduced free-energy based coarse graining approach - the Conditional Reversible Work (CRW) method[13] - in which the interaction potentials are obtained *at pair level*. Hence, we obtain a genuine pair potential. This potential is physical because it describes the conditional interaction free energy between two groups of atoms as a function of the distance separating them. The imposed condition is that the two atom groups are considered within their covalently bonded chemical environments which effectively limits the configuration space for interaction of the two groups in the CRW calculation. In this work, we investigate different free energy calculation methods for obtaining the CRW-CG potentials for hexane and study the chemical transferability of these potentials to describe liquid phase properties of longer alkanes. The thermodynamic transferability of the models is tested by examining the liquid phase properties in a range of different temperatures.

4.2 Coarse Graining using the Conditional Reversible Work Method

The CRW pair potential is a conditional interaction free energy between the atoms represented by two nonbonded beads of the CG model. The condition refers to the free energy being calculated in the presence of chemical groups belonging to the immediate (bonded) environment of the beads. This condition reduces the available configuration space of the groups for which the CRW potential is computed, because only the relative orientations of these groups, sterically allowed by surrounding chemical groups, are sampled. CRW potentials can be obtained from free energy calculations in a vacuum environment, or alternatively in a condensed-phase liquid environment, using different free energy sampling methods. Two methods are explored in this paper. In the first method, previously published by us,[13] the CRW potentials are calculated employing a thermodynamic cycle based on two potential of mean force calculations. This method is computationally cheap when the two molecules are sampled in vacuum, but it becomes expensive and difficult to converge when the molecules are part of a bulk liquid simulation. The new second method that we introduce and test in this paper is based on thermodynamic perturbation theory and takes advantage of the full ensemble of mutual configurations of

all possible molecule pairs in the pure component liquid phase. The two methods are explained below.

4.2.1 Thermodynamic Cycle: Potential of Mean Force Method

The thermodynamic cycle presented in Fig. 4.1 illustrates the calculation of the CRW pair potential $U_{eff}(r)$. Hexane is taken as an example, illustrating the calculation of the effective pair potential between two beads that are mapped on the two central methylene groups of the hexane molecules. In the process denoted by the vertical arrow, direct interactions are introduced between the mapped atom groups (beads) at a center of mass distance r; the free energy of this coupling process defines the CRW potential $U_{eff}(r)$. One way to obtain the CRW potentials is to compute two potentials of mean force. $RW(r)$ represents the potential of mean force of the CG degree of freedom r. $RW_{excl}(r)$ represents the same quantity, obtained by excluding the direct interactions between the mapped atom groups. Using the thermodynamic cycle and using $U_{eff}(\infty) = 0$ it follows that $U_{eff}(r) = RW(r) - RW_{excl}(r)$. The calculation of $U_{eff}(r)$ in a vacuum environment may be performed with constraint dynamics in which all degrees of freedom other than r are sampled and the average constraint force is integrated to obtain $RW(r)$ and $RW_{excl}(r)$. The models obtained in this way will be referred to as CG_{vac}. The same procedure applied in the liquid phase of the pure component leads to models that will be referred to as CG_{liq}.[13]

4.2.2 Direct Calculation by Free Energy Perturbation

CG_{vac} models are readily computed using the potential of mean force calculations discussed above. It is however not a priori clear if these models are transferable to condensed phase systems at liquid-like densities, therefore CG_{liq} models are calculated as well. The calculation of $RW_{excl}(r)$ in the liquid phase is computationally expensive and is prone to larger statistical uncertainties than $RW(r)$, which, in the liquid phase, can be obtained by inverting the radial distribution function obtained from the mapped CG coordinates. Averaging over all pairs improves the statistical accuracy of $RW(r)$ over $RW_{excl}(r)$, which is calculated by considering a single pair only. An alternative calculation route that takes full advantage of the sampled distance distribution is based on free energy perturbation[14] and provides a direct calculation of $U_{eff}(r)$. In this method, the reference ensemble corresponds to config-

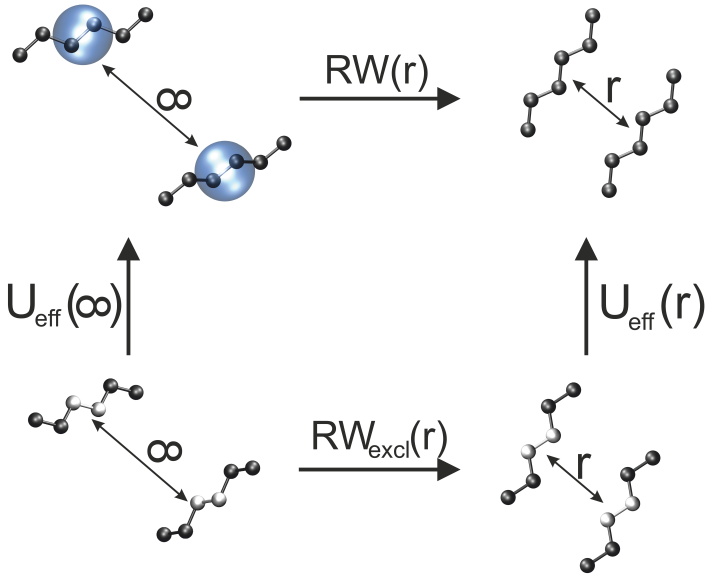


Figure 4.1: Thermodynamic cycle used to calculate the CRW pair potential $U_{eff}(r)$ between two beads mapped on the two central methylene groups of two hexane molecules. $U_{eff}(r)$ is obtained from the difference between two potentials of mean force: $U_{eff}(r) = RW(r) - RW_{excl}(r)$. While $RW(r)$ includes the contributions of all interactions between the atoms of the two fine-grained molecules, $RW_{excl}(r)$ excludes direct interactions between the atoms mapped into the CG beads.

urations of a liquid state trajectory generated by the full (unperturbed) Hamiltonian of the fine-grained system. The perturbation corresponds to selecting at random two molecules of the reference ensemble and "switching off" the direct nonbonded interaction between the mapped atom groups on these molecules. The corresponding energy depends on r as well as on the mutual orientation of the two molecules and is denoted ΔU. The probability density distribution of the perturbation energy in the reference ensemble is denoted $P(\Delta U; r)$ and relates to the CRW pair potential according to

$$U_{eff}(r) = \frac{1}{\beta} \ln \int_{-\infty}^{\infty} e^{-\beta \Delta U} P(\Delta U; r) d\Delta U \tag{4.1}$$

where $\beta = (k_B T)^{-1}$ with k_B representing the Boltzmann constant and T the temperature. The CRW models obtained with this method will be referred to as CG_{per}.

4.3 Models

The CRW-CG models presented in this paper have been developed using the OPLS forcefield (FF) [15] as the fine-grained model. The OPLS-FF uses a united-atom (UA) description of the methyl and methylene units of alkane chains, *i.e.* the aliphatic hydrogen atoms are not explicitly represented. The representability and transferability of the CRW models will be discussed by comparison against the thermodynamic data predicted by the fine-grained model.

The CRW models developed in this work all map two united atoms in a single interaction site. Considering that we deal with linear alkane chains we can identify two different kind of beads: the E beads that represent the head and tail groups of the alkane chain ($CH_3 - CH_2-$) and the I bead that represents the internal "monomers" of the chain ($-CH_2 - CH_2-$). As shown in Fig. 4.2, combining these two beads allows to describe linear alkane chains with an even number of carbon atoms. In this work we study even alkane chains from butane to dodecane. Three sets of CRW-based nonbonded potentials have been employed in this work. The first two potentials (CG_{vac} and CG_{liq}) are based on the thermodynamic cycle method, while the third potential (CG_{per}) is developed based on the free energy perturbation method.

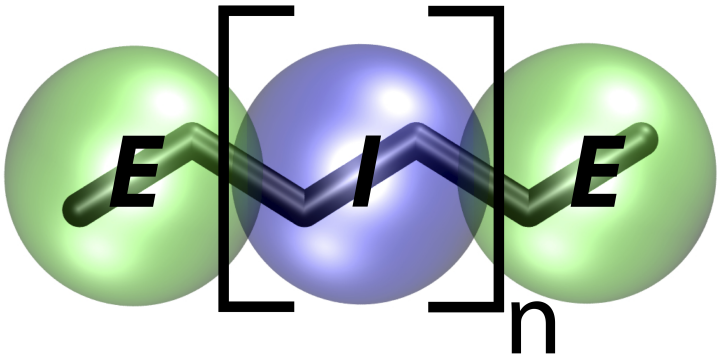


Figure 4.2: Mapping scheme of the CG alkane model. Beads E represent the terminal ethyl groups of the chain, beads I represent the $-CH_2-CH_2-$ repeating units. In this paper we study chains with $n = 0, 1, 2, 3, 4$.

4.4 Computational Details

In order to determine the quality of the different CRW-CG models we compare their predictions of thermodynamic properties with the results obtained using the fine-grained UA model. The representability is investigated by comparing the den-

sity, the liquid structure, liquid surface tension, heat capacity, thermal expansion coefficient, isothermal compressibility and the excess free energy of liquid hexane (relative to an ideal gas at the same density) at 298.15 K and 1 atm. The thermodynamic transferability is investigated by comparing the liquid densities obtained with the models at different temperatures, spanning over almost the entire hexane liquid range at 1 atm. The chemical transferability is investigated by comparing the density as a function of the temperature for the series of even alkanes in the liquid region of the phase diagram at 1 atm covering a temperature range between $T_{melting} + 10$ K and $T_{boiling} - 10$ K for all systems. The simulations have been performed using the GROMACS simulation package[16].

4.4.1 MD simulations with the OPLS model

All fine-grained MD simulations of pure alkanes have been performed with a rectangular periodic box containing 512 alkane molecules. Simulation trajectories up to 10 ns were accumulated using the stochastic dynamics (SD) integrator of GROMACS with a time-step of 1 fs and an inverse friction coefficient of 0.2 ps. The pressure of the simulation box was maintained at 1 atm using a Parrinello-Rahman barostat[17] with a coupling time of 1 ps. Lennard-Jones interactions were evaluated using a twin range cut-off scheme with cut-off distances of 1.0 nm and 1.3 nm. Nonbonded interactions at distances smaller than 1.0 nm were updated every time step while the interactions in the range between 1.0 and 1.3 nm were updated every 10th time step. The densities were calculated by averaging over the last 8 ns in each simulation. In order to calculate the heat capacity and the isothermal compressibility of hexane the fluctuations of volume and enthalpy have been analyzed in an extended NpT simulation of 50 ns at 298.15 K. The surface tension was calculated from a 50 ns NVT simulation of an half empty box obtained extending the length of an equilibrated cubic box of liquid. The ideal gas heat capacity of hexane was calculated from a 50 ns NVT simulation of 512 hexane molecules in a large box, neglecting the nonbonded interactions between different molecules. The excess free energy of liquid hexane was calculated using thermodynamic integration (TI)[18]. A series of 20 simulations associated with twenty equally spaced values of the coupling parameter λ have been performed to calculate the free energy of the process of switching on the direct intermolecular interactions between 512 hexane molecules. Every λ-point was simulated for 1 ns at NVT conditions, with the density fixed to the average density of liquid hexane as obtained from the UA model, and the last 0.8 ns were

used for data collection. Soft core potentials have been employed to avoid singularities at small values of λ.[19] All the other simulation parameters were identical to the ones used for the determination of the density of pure alkane. Properties of a mixture of hexane and dodecane were calculated for mole fraction compositions $x_{dodecane} = 0.25, 0.33, 0.5, 0.66, 0.75$. The simulation parameters for this series of simulations were identical to those employed for the determination of the density of the pure alkanes.

4.4.2 CG potentials

The CG bonded interactions were developed by sampling the conformation space of a OPLS-UA dodecane chain in vacuum in a 10 ns MD simulation with a time step of 1 fs. This simulation used a stochastic dynamic integrator with an inverse friction coefficient of 0.2 ps. The VOTCA[20] package has been employed to calculate the CG bond length, bond angle and torsion distributions and invert them in order to obtain potentials for bonds, angles and dihedrals for the CG model.
The first two sets of nonbonded CRW-CG interaction potentials (CG_{vac} and CG_{liq}) have been developed in a previous work, and we refer to it for the details of their development[13]. The third set of CG interaction potentials (CG_{per}) has been developed using the thermodynamic perturbation theory (see section 4.2.2). The last 8 ns of the simulation of liquid hexane at 298.15 K and 1 atm have been used to develop the nonbonded potentials using Eq.4.1. The considered range of the CRW pair potentials goes from 0.32 nm to 1.1 nm with intervals of 0.02 nm. The CRW pair potential was shifted in order to be 0 at a distance of 1.1 nm, beyond this distance the direct interaction is neglected. The repulsive part of the potential at short distance was calculated through spline interpolation. All the CG interaction potentials were used as tabulated potentials in all the CG simulations.

4.4.3 MD simulations with the CG models

The CG MD simulations were performed with system sizes, simulation times and simulation settings identical to those described in section 4.4.1. The only differences are the time step, which was chosen 2 fs in the CG simulations, and the interaction cut-off, which was chosen 1.1 nm. Furthermore, no soft core potentials were necessary to avoid singularities in the TI calculations, since the CG interaction potentials are interpolated at small distance and do not diverge. The TI calculations

were conducted under NVT conditions at the density (1 atm) predicted by the CG model.

4.5 Results and Discussion

4.5.1 Non bonded interaction potentials

Fig. 4.3 shows the three different sets of CG interaction potentials obtained from CRW calculations. We define the cutoff of the CG models to be 1.1 nm, therefore all the interaction potentials should go to zero at this distance. For the models CG_{vac} and CG_{liq} this is achieved by integrating the constraint forces with this distance as integration limit,[13] while for the CG_{per} model the interaction energy is shifted to be zero at the cut-off distance.

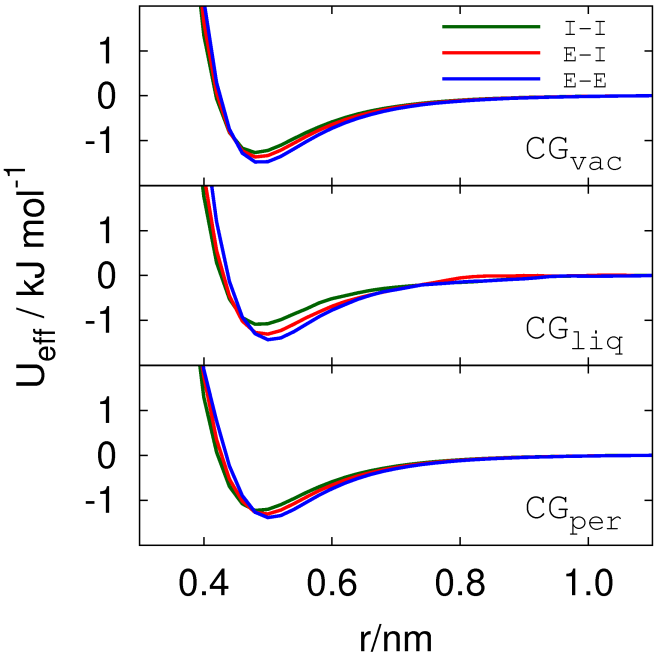


Figure 4.3: CRW potentials used in this work. CG_{vac} and CG_{liq} have been developed using the thermodynamic cycle method (298.15 K) with two molecules in vacuum and in liquid hexane environments, respectively. The set of interaction potentials CG_{per} has been calculated by applying the free energy perturbation method. The potentials of the CG_{liq} model are more noisy than the potentials of the other two models due to poorer statistics in sampling of the average constraint force.

The CRW potentials corresponding to the three models CG_{vac}, CG_{liq} and CG_{per} are qualitatively similar. The potentials all have a monotonically varying tail, as ex-

pected for the long range interaction between two groups of atoms. The interaction potentials characterizing the $E-E$ interaction are deeper and have a bigger excluded volume compared with the potentials that represent the $I-I$ interaction. This reflects the presence of a CH_3 group in the E beads, that makes the interaction site bulkier and more attractive. Also in all three sets, the $E-I$ interaction potential is located in between the other two interaction potentials, confirming that it represents a "mixed" interaction.

The CG_{liq} potentials contain small irregularities in the tails. This is due to difficulties of fully converging the constraint force when two molecule are far away in the liquid phase. The long range decay of the interaction is similar for the CG_{vac}, CG_{liq} and CG_{per} potentials. The similarity between the three sets is not surprising, since in an isotropic liquid the relative orientations of two molecules are not correlated at large distances, therefore there is no difference between the tails of the potentials obtained by sampling in a vacuum and in a liquid environment. If we however consider the repulsive part of the potential at short distances, we see that there are differences between the sampling in vacuum (CG_{vac}) and in the liquid phase (CG_{liq} and CG_{per}). The short range repulsive regions of the CG_{liq} and CG_{per} potentials are almost identical but deviate slightly from the CG_{vac} model. This difference originates from subtle differences in the short range packing of molecules in the liquid phase and dimers in the gas phase. The CG_{vac} model has a steeper short range repulsion compared to the other models.

4.5.2 Representability and thermodynamic transferability

We investigated several thermodynamic properties of liquid hexane at the state point where the CRW hexane models were developed; *i.e.* 298.15 K and 1 atm for the CRW models obtained from simulations of the liquid phase. We also tested the CG_{vac} model at the same conditions in order to compare the predictions of this model with the models developed in the liquid phase. Table 4.1 summarizes the thermodynamic quantities considered. Note that the investigated quantities include thermodynamic response functions (heat capacity, thermal expansion coefficient and isothermal compressibility), which are related to equilibrium fluctuations that determine state point transferability. Since these fluctuations depend on the nature of the potentials which is related to the degree of coarse graining, state point transferability is expected to be limited. Here, it will be investigated how far the current CRW

models reach in achieving temperature transferability in the liquid phase.

	Hexane						Dodecane
	d $\left[\frac{kg}{m^3}\right]$	ΔF_{ex} $\left[\frac{kJ}{mol}\right]$	γ $\left[\frac{mN}{m}\right]$	c_p $\left[\frac{J}{mol\,K}\right]$	κ_T $\left[\frac{10^{-9}}{Pa}\right]$	α_p $\left[\frac{10^{-3}}{K}\right]$	α_p $\left[\frac{10^{-3}}{K}\right]$
UA	667.7	-15.5	17.6	190.5	1.7	1.31	1.11
CG_{vac}	727.9	-6.6	18.5	96.9	1.8	1.16	0.94
CG_{liq}	668.7	-8.0	14.2	96.4	2.4	1.16	0.97
CG_{per}	706.9	-11.1	16.4	92.7	2.2	1.06	0.85
$CG_{per}(sh)$	659.1	-7.9	9.9	95.4	3.6	1.35	1.09

Table 4.1: Predicted thermodynamic properties of liquid hexane and liquid dodecane at 298.15 K and 1 atm for the UA model, the three CRW-CG models (CG_{vac}, CG_{liq}, CG_{per}) and the potential shifted in order to reproduce the hexane UA density ($CG_{per}(sh)$).

The first investigated obsevable is the liquid **mass density** (d). The CG_{liq} model reproduces the liquid hexane density of the fine-grained UA model closest (Table 4.1). Comparing the result of the CG_{liq} model with the results we have previously published[13], where a cut-off of 1.3 nm was employed, we can see that by reducing the cut-off the density increases. This is somehow counterintuitive, and can only be explained accounting for compensation of errors due to the noise on the CG_{liq} potentials. The CG_{vac} model predicts a density which is 9% larger than the UA density, while the CG_{per} model predicts a density that is too large by 6%. We tried to improve the ability of the CG_{per} model to reproduce the density of the fine-grained UA model by shifting the cut-off distance. We find that with a cut off distance of 0.92 nm the density can be reproduced within an error of the 1% (the corresponding model is referred to as $CG_{per}(sh)$).

The **excess free energy** (ΔF_{ex}) of liquid hexane could not be quantitatively reproduced by any of the CG models. We however point out that very small changes in the pair potential can lead to large variations of a few kJ mol^{-1} in the predicted excess free energies.[21] The CG_{per} model yields the most accurate representation of ΔF_{ex}; shifting the potential in order to reproduce the density ($CG_{per}(sh)$) leads to a poorer description of ΔF_{ex}.

The **surface tension** (γ) has been obtained from a simulation of a periodic box with a slab of hexane periodic in x and y-directions, using

$$\gamma = \frac{L_z}{2}\left\langle p_{zz} - \frac{p_{xx}+p_{yy}}{2}\right\rangle \tag{4.2}$$

where p_{xx}, p_{yy} and p_{zz} are the diagonal components of the pressure tensor, and L_z is the linear dimension of the box in the z-direction. Table 4.1 shows that the CRW models reproduce the surface tension in close agreement with the UA model, except for model $CG_{per}(sh)$. This is imputable to the shorter cut-off distance employed for this model which causes a reduced cohesion between the molecules and a decrease in the energetic cost of creating an interface.

The **molar heat capacity** (c_p) has been calculated using the enthalpy fluctuation formula

$$c_p = \frac{\langle H^2\rangle - \langle H\rangle^2}{nk_BT^2} \tag{4.3}$$

where the numerator denotes the variance of the enthalpy ($H = E + pV$), n the number of molecules (in moles), and the brackets denote averages in the constant pressure-temperature ensemble. The heat capacities reported in table 4.1 show small variations among the CRW models while the fine-grained UA model has a significantly larger heat capacity. This discrepancy is largely owing to the smaller number of internal degrees of freedom of the CG hexane models. The internal contributions to the heat capacities were determined using calculations in the gas phase and amount to 137 J mol^{-1} K^{-1} for the UA model and 56 J mol^{-1} K^{-1} for the CG models. After removing the internal parts, the excess heat capacities of the CRW models are up to 30 % lower in comparison to the UA model. This underestimation of the heat capacity reflects the effect of coarse graining the intermolecular interactions which invariably leads to flattening out of the potential energy surface (PES) and smaller amplitudes of the energy fluctuations.

The **isothermal compressibility** (κ_T) of the model has been obtained by analyzing the volume fluctuations in a NpT simulation at 298.15 K and 1 atm using the fluctuation formula

$$\kappa_T = \frac{\langle V^2\rangle - \langle V\rangle^2}{\langle V\rangle k_BT^2} \tag{4.4}$$

Comparison between the UA and CRW models shows that κ_T is overestimated for the three CRW models developed in the liquid phase. The CG_{vac} model also overestimates the compressibility but to a significantly smaller extent compared with the other CRW models. While flattening of the PES causes reduced energy fluctuations and smaller heat capacity, it causes enhanced volume fluctuations and larger compressibility.

The **thermal expansion coefficient** (α_p) has been calculated from the linear fit of the density d vs. the temperature T (see Fig.4.6). The temperature range shown in Fig.4.6 spans almost the entire liquid range of hexane (shown with the black symbols). The $CG_{per}(sh)$ model reproduces α_p in close agreement with the fine-grained UA model while the other CRW models tend to slightly underestimate α_p. The response of the density to a variation in temperature is determined by equilibrium fluctuations of the volume (V) and enthalpy (H) of the NpT system; *i.e.* $\alpha_p = \langle \delta V \delta H \rangle / k_B T^2 \langle V \rangle$, where δV and δH are the volume and enthalpy fluctuations, respectively. We thus see that a reasonably good representation of the thermal expansion coefficients of the CRW models can be obtained despite smaller enthalpy fluctuations (heat capacity) and larger volume fluctuations (compressibility). While systematic coarse graining invariably leads to softer interactions and flattening of the PES, good temperature transferability of the density can therefore be obtained with the CRW models.

To further examine the representability of the models we examined the **radial distribution functions** (RDFs) between the hexane centers of mass and the different types of CG beads obtained with the UA and CRW models in Fig.4.4. Small discrepancies are observed between the UA and CRW RDFs when considering the COM and $I-I$ correlations. However, the overall agreement between the RDFs obtained with the UA and the CRW models is very good.

4.5.3 Chemical transferability

The chemical transferability of the CRW-CG models is a key quantity, which, for the system considered in this work, determines whether the CRW potentials of the E-type and I-type beads can be used in CG simulations of different aliphatic hydrocarbons. The first test of chemical transferability involves the prediction of the density of hexane, octane, decane and dodecane at 1 atm and 298.15 K. We exclude butane from this series, since 298.15 K is above the boiling point of this compound. The densities are reported in Fig. 4.5. All models are able to qualitatively reproduce the

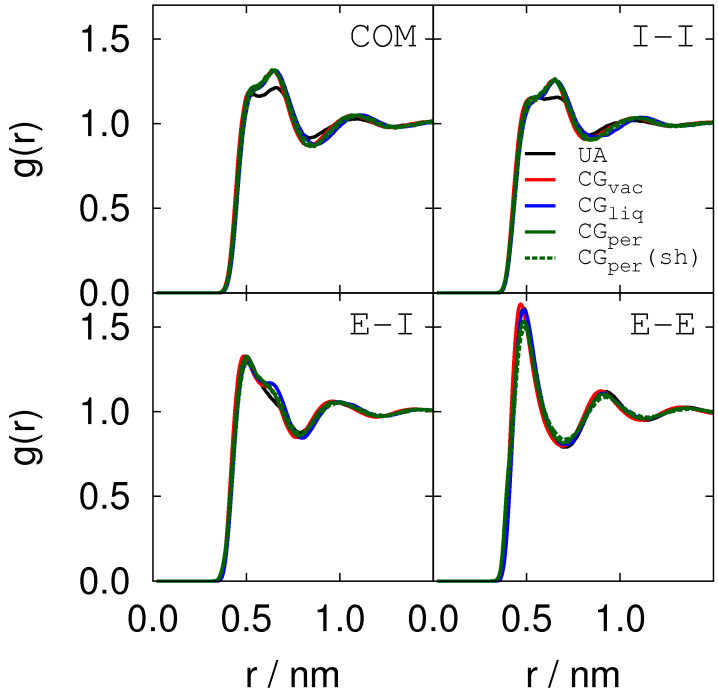


Figure 4.4: Radial distribution functions for the hexane centers of mass (COM) and the different bead types. Sampling was performed over the last 8 ns of a NpT simulation of liquid hexane at 298.15 K and 1 atm employing the different CRW models.

dependence of the density on the chain length of the alkane. The observed curvature can be ascribed to the free volume of the end groups and is reproduced with all CRW-CG models. Fig. 4.6 compares the densities obtained with the fine grained UA and CRW-CG models of the different alkanes in the liquid temperature range. All CRW-CG models represent the temperature dependence qualitatively very well. The $CG_{liq}(sh)$ model shows the best quantitative agreement with the fine-grained UA model for all alkane systems.

Fig. 4.7 shows the COM RDFs of liquid dodecane and also the RDFs between the different bead types. The CRW-CG models are able to predict the positions of the various peaks in agreement with the fine-grained dodecane model, but tend to underestimate their amplitudes in particular for the COM RDFs and the RDFs between the I beads. The correlations involving the E-beads are better reproduced.

Finally, we have investigated binary mixtures of hexane and dodecane. All CRW-CG models reproduce the density of the mixture within an error of 9% (not shown). Fig. 4.8 presents the excess volume of mixing obtained with the different models. The best agreement with the fine-grained model is obtained with CG_{vac}, which is

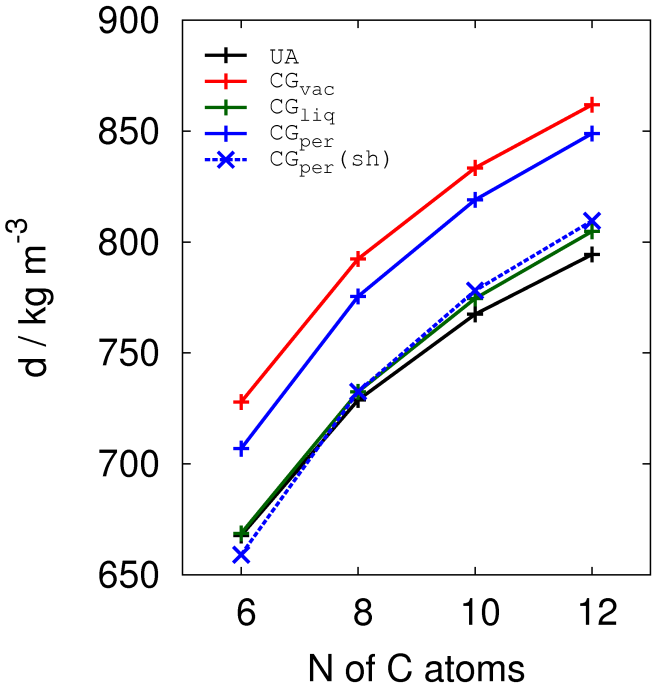


Figure 4.5: Mass densities of n-alkanes obtained with the CRW-CG models and the fine-grained UA model for different alkane chain lengths (298.15 K and 1 atm.)

also in good agreement with available experimental data [22] for this system. Clearly, the $CG_{per}(sh)$ model overestimates the excess volume, indicating that modifying the potential in order to fit one property (the density of pure liquid hexane), deteriorates the representability with respect to other properties (which further include the excess free energy and surface tension of hexane shown in Table 4.1). Fig. 4.9 shows the center of mass RDFs, $g_{ij}(r)$, of a 50% hexane/dodecane mixture (upper panel) and the Kirkwood-Buff integrals (KBIs) [23] as a function of mixture composition (lower panel). The KBIs measure the preferential solvation of mixture components [24] and are related to the composition dependence of the chemical potentials. The KBIs were calculated according to

$$G_{ij} = \lim_{R\to\infty} 4\pi \int_0^R [g_{ij}(r) - 1] r^2 dr \tag{4.5}$$

While still oscillating, the integral in Eq. 4.5 converges in the distance range $1.5 < R < 2.0$ nm. The KBIs shown in Fig. 4.9 correspond to averages taken in this range.

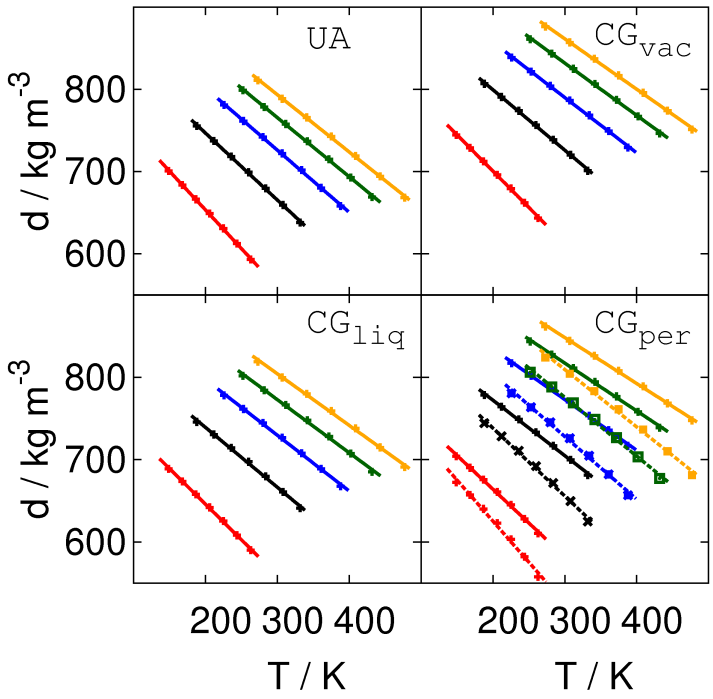


Figure 4.6: Mass densities of butane (red), hexane (black), octane (blue), decane (green) and dodecane (orange) calculated at 1 atm over the entire temperature range of the liquid phase with the UA and four different CRW-CG models. In the last panel, the dashed lines show the densities obtained with the $CG_{liq}(sh)$ model. The lines are the linear fits of the data points.

Within the uncertainty of the error bars, the CRW-CG models reproduce the KBIs obtained with the fine-grained model.

4.6 Discussion and Conclusions

In this paper we have applied the conditional reversible work (CRW) method [13] to develop coarse-grained pair potentials for noncovalent interactions in liquid hydrocarbons. The CRW pair potential satisfies two criteria that are important to achieve chemical transferability. First, the potential is physical, *i.e.* it describes the interaction free energy between the two groups of atoms composing the coarse-grained particles as a function of the center of mass distance between them. Second, the CRW pair potential is calculated under the imposed condition that the interacting atom groups are considered in the chemical environment of the molecule they are part of. In this paper, CRW potentials were developed for the CH_3-CH_2- and $-CH_2-CH_2-$ interaction sites in hexane, and have been used to study butane, hexane, octane, decane and dodecane in the liquid phase at different temperatures and 1 atm pressure. We

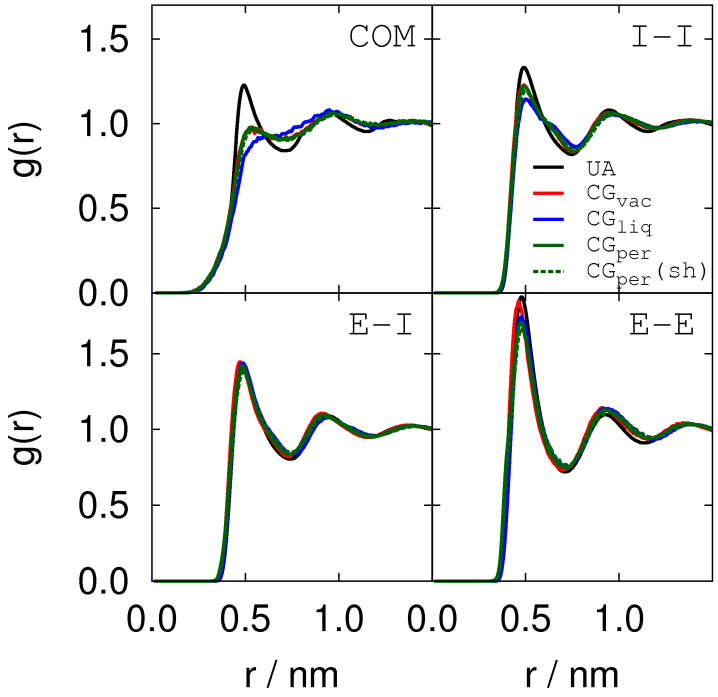


Figure 4.7: Dodecane RDFs calculated between the COM of the molecules and between the different bead types during the last 8 ns of a NPT simulation of liquid dodecane at 298.15 K and 1 atm employing the different CRW-CG models.

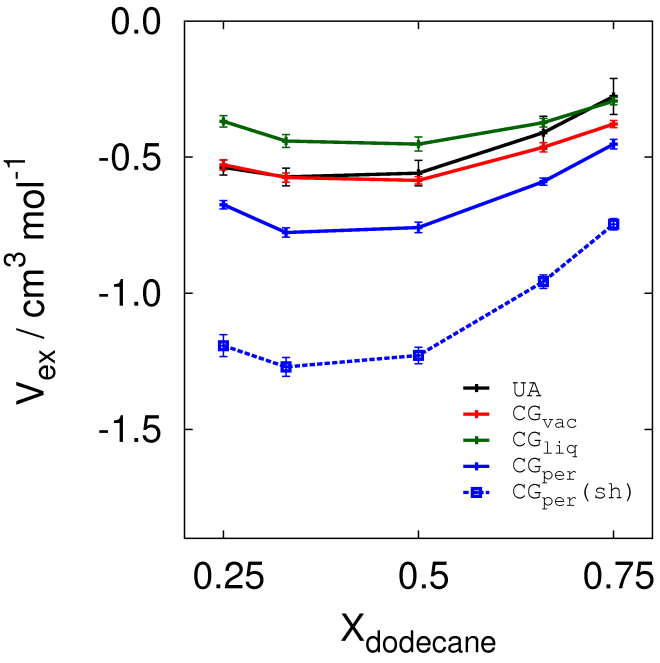


Figure 4.8: Excess volume of mixing of hexane and dodecane as a function of mixture composition (298.15 K, 1 atm.).

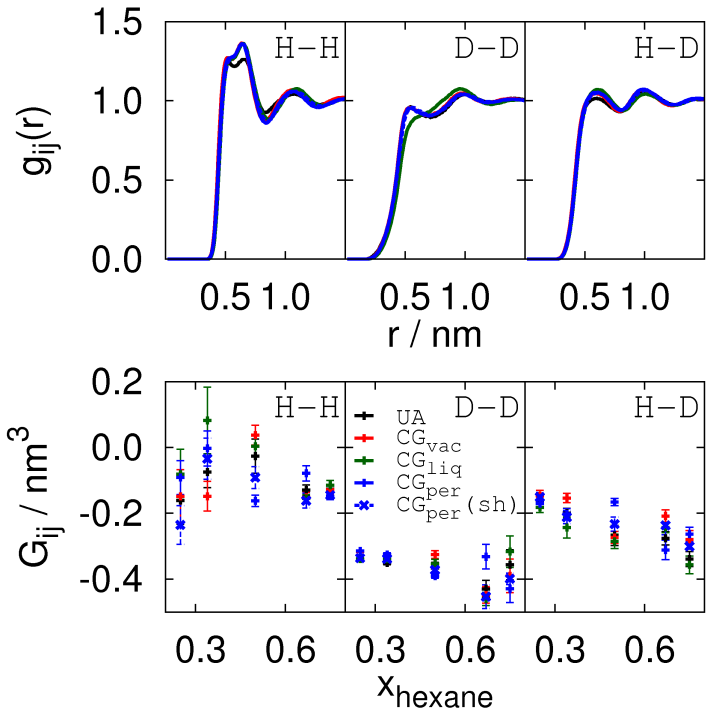


Figure 4.9: Center of mass RDFs of the mixture components hexane (H) and dodecane (D) in a 50% hexane-dodecane mixture (upper panel) and Kirkwood-Buff integrals of the mixture components as a function of mixture composition (lower panel). All data at 298.15 K and 1 atm.

find that the CRW alkane model exhibits reasonable thermodynamic transferability with a good description of the mass density (1 atm) at all temperatures in the liquid region of the phase diagram. The potentials are furthermore chemically transferable to systems with different aliphatic chains lengths as well as to binary mixtures of hexane and dodecane. Several CRW models were developed with different free-energy calculation methods. The free energy perturbation method provides a fast and facile route to obtain the CRW coarse-grained potentials from a single atomistic simulation of liquid hexane.

Clearly, the assumption of pairwise additivity of the N-body coarse-grained potential limits the applicability of the derived CRW models to specific classes of systems while alternative methods, which automatically include average multi-body effects in a set of *effective* pair potentials, are available.[1–11] These effective pair potentials, however, cannot always be considered to describe a physical interaction, but should rather be interpreted as numerical constructs optimized to reproduce preselected target quantities[25]. The CRW method depicted in Fig. 4.1 can however in a straightforward way be extended, for example, to develop CG implicit-solvent models for macromolecules such as polyelectrolytes in aqueous solution[26]. Com-

bined with earlier approaches to account for multibody effects[27,28] the advantage is kept over alternative coarse graining approaches that the physical behaviour of the system can be related to the nature of the interactions, which are physical.

4.7 Acknowledgement

The authors thank Reinier L.C. Akkermans for insightful discussions.

4.8 Bibliography

[1] F. Ercolessi and J. Adams. *Europhys. Lett.*, 26:583, 1994.

[2] A. Lyubartsev and A. Laaksonen. *Phys. Rev. E*, 52:3730, 1995.

[3] D. Reith, M. Pütz, and F. Müller-Plathe. *J. Comput. Chem.*, 13:1624, 2003.

[4] S. Izvekov and G. A. Voth. *J. Phys. Chem. B*, 109:2469, 2005.

[5] S. J. Marrink, H. R. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *J. Phys. Chem. B*, 111:7812–7824, 2007.

[6] M. S. Shell. *J. Chem. Phys.*, 129:144108, 2008.

[7] W. G. Noid, J. Chu, G. S. Ayton, V. K., S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen. *J. Chem. Phys.*, 128:244114, 2008.

[8] R. L. C. Akkermans. *J. Chem. Phys.*, 128:244904, 2008.

[9] Q. Wang, D. J. Keffer, D. M. Nicholson, and J. B. Thomas. *Phys. Rev. E*, 81:061204, 2010.

[10] P. Ganguly, D. Mukherji, C. Junghans, and N. F. A. van der Vegt. *J. Chem. Theory Comput.*, 8:1802, 2012.

[11] J. R. Allison, S. Riniker, and W. F. van Gunsteren. *J. Chem. Phys.*, 136:054505, 2012.

[12] D. Fritz, K. Koschke, V. A. Harmandaris, N. F. A. van der Vegt, and K. Kremer. *Phys. Chem. Chem. Phys.*, 13:10412, 2011.

[13] E. Brini, V. Marcon, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 13:10468, 2011.

[14] C. Chipot and A. Pohorille. *Free Energy Calculations: Theory and applications in chemistry and biology*. Springer, 1st edition, 2007.

[15] W. L. Jorgensen and J. Tirado-Rives. *J. Am. Chem. Soc.*, 110:1657, 1988.

[16] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Theory Comput.*, 4:435, 2008.

[17] M. Parrinello and A. Rahman. *J Appl. Phys.*, 12:7182, 1981.

[18] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Academic Press, 2nd edition, 2001.

[19] B. Hess and N. F. A. van der Vegt. *J. Phys. Chem. B*, 110:17616, 2006.

[20] V. Ruehle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko. *J. Chem. Theory Comput.*, 12:3211, 2009.

[21] E. Brini, C. R. Herbers, G. Deichmann, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, page DOI: 10.1039/C2CP40735C, 2012.

[22] J. D. Gomez-Ibanez and C. T. Liu. *J. Phys. Chem.*, 67:1388, 1963.

[23] J. G. Kirkwood and F. P. Buff. *J. Chem. Phys.*, 19:774, 1951.

[24] A. Ben-Naim. *Molecular Theory of Solutions*. Oxford University Press, 2006.

[25] A. A. Louis. *J. Phys. Cond. Mat.*, 14:9187, 2002.

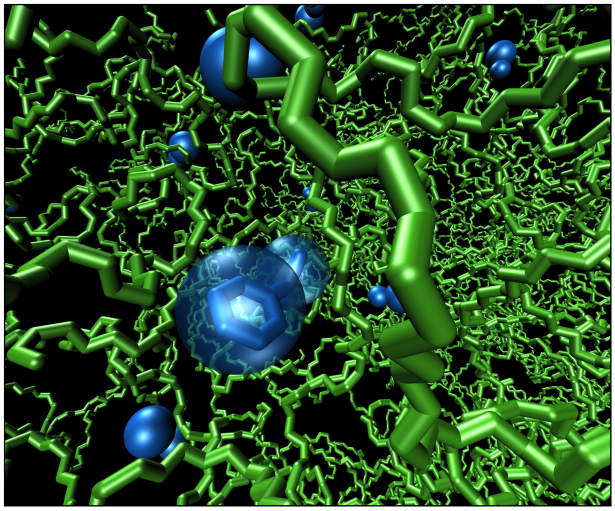
[26] C. Li, J. Shen, C. Peter, and N. F. A. van der Vegt. *Macromolecules*, 45:2551, 2012.

[27] B. Hess, C. Holm, and N. F. A. van der Vegt. *Phys. Rev. Lett.*, 96:147801, 2006.

[28] J.W. Shen, L. Li, N. F. A. van der Vegt, and C. Peter. *J. Chem. Theory Comput.*, 7:1916, 2011.

5 Thermodynamic transferability of coarse-grained potentials for polymer-additive systems

In this work we study the transferability of systematically coarse-grained (CG) potentials for polymer-additive systems. The CG nonbonded potentials between the polymer (atactic polystyrene) and three different additives (ethylbenzene, methane and neopentane) are derived using the Conditional Reversible Work (CRW) method, recently proposed by us [Brini et al., *Phys. Chem. Chem. Phys.*, 2011, **13**, 10468]. A CRW-based effective pair potential corresponds to the interaction free energy between the two atom groups of an atomistic parent model that represent the coarse-grained interaction sites. Since the CRW coarse-graining procedure does not involve any form of parameterisation, thermodynamic and structural properties of the condensed phase are predictions of the model. We show in this work that CRW-based CG models of polymer-additive systems are capable of predicting the correct structural correlations in the mixture. Furthermore, the excess chemical potentials of the additives obtained with the CRW-based CG models and the united-atom parent models are in satisfactory agreement and the CRW-based CG models show a good temperature transferability. The temperature transferability of the model is discussed by analysing the entropic and enthalpic contributions to the excess chemical potentials. We find that CRW-based CG models provide good predictions of the excess entropies, while discrepancies are observed in the excess enthalpies. Overall, we show that the CRW CG potentials are suitable to model structural and thermodynamic properties of polymer-penetrant systems.



5.1 Introduction

A detailed microscopic picture is necessary to understand molecular transport processes in polymer-penetrant systems, like liquid and vapour separation membranes, plasticizers in engineering thermoplastics or polymer dissolution processes. Computer simulations have proven to be a useful tool to study the permeation of additives in polymer matrices.[1,2] Atomistic force fields (FFs) are in principle capable of describing these processes accurately, however, often they can not reach the time and length scale needed to simulate, e.g., a polymer swelling or dissolution process. In order to extend time and length scales, coarse-grained (CG) models can be used. To keep a link to the specific chemistry of the system of interest, systematically coarse-grained models are needed, which have been developed in recent years for various macromolecular systems.[3–11] A CG model is usually obtained by selecting a mapping scheme which merges neighbouring atoms into so-called "super-atoms" or CG beads and subsequently applying a systematic coarse-graining procedure that provides the corresponding bonded and nonbonded interaction potentials. Reducing the number of particles in the system leads to a reduction of degrees of freedom (DOF), enabling the possibility of studying phenomena happening on time and length scales that atomistic simulations can not capture.

In order to study specific polymer-penetrant systems, the CG potentials should be capable of reproducing both thermodynamic and structural properties of the system. The method used to develop these potentials must therefore cast a certain amount of chemical information into the CG potentials, either through a parameterisation procedure or by other means. The interaction potentials can be parametrised in order to reproduce thermodynamic properties (e.g. partitioning free energy between a polar and an apolar phase[12], liquid-vapour equilibria[13], partition function in the gas-phase[14] or equation of state[15]) or microscopic properties (e.g. liquid structure[16,17] or force distribution[18]). These approaches provide effective pair potentials that are able to reproduce the target property, but nothing can a priori be said about their ability to reproduce other properties or their ability of predicting properties at a state point different from the one used in the parametrisation process. A completely different approach to developing coarse-grained FFs is based on applying the CG mapping scheme at the atomistic level and calculating the effective interaction between the mapped atom groups by a suitable averaging procedure. Examples of quanties that can be used as effective potentials are: the pair potential of

mean force [19–21], the effective force [22] and the conditional reversible work (interaction free energy) [23]. This class of methods provides pair potentials that carry-over chemistry-specific information from the detailed atomistic level of description to a coarse-grained mesoscopic level in a systematic manner without resorting to parameterisation. In comparison with potentials obtained by a parameterisation procedure, these methods yield pair potentials which can be more easily related to the distance-dependent interaction between the coarse-grained atom groups. Of course, also with these methods pairwise addivity of the coarse-grained potential is assumed and nothing can a priori be said about the effectiveness of the potentials to reproduce any property. In this work, the CG non-bonded interaction potentials are developed using the Conditional Reversible Work (CRW) method [23], which was previously shown to model the liquid structure and the density of molecular liquids in good agreement with atomistic simulations. Here, the interaction free energy between the groups of atoms that the beads represent is used as an effective pair potential. An advantage of this method is that effective pair potentials developed in vacuum can be applied in the condensed phase, therefore these potentials are cheap to obtain. [4,23].

The quality of CG models is often discussed in terms of representability and transferability. The representability is the ability of the CG model to predict properties at the state point used for its parametrisation. The transferability is the ability of the model to predict properties at different state points. In our work, the pairwise interactions are developed in vacuum (see section 5.4.1) and are applied in the condensed phase. Therefore, only the transferability of our model will be investigated. In order to understand the quality of the CRW potentials developed in this work, we focus our attention on the prediction of the excess chemical potential (ECP) of small molecules (ethylbenzene (EB), methane (ME) and neopentane (NP)) in a melt of atatic polystyrene (PS) [4]. The choice to study ECPs has several advantages. First of all, this property can be calculated both on an atomistic and CG level (see section 5.2.2), offering the possibility to compare the results obtained using the atomistic model and the derived CG model. This comparison is important since the aim of a systematically developed CG model is to reproduce the results obtained by the parent atomistic model from which it is derived. The second advantage of investigating ECPs is the fact that this property is extremely sensitive to the quality of the pair potentials, therefore providing a good means to investigate the quality of our pair potential. In order to better understand why CG models are transferable we will study also the entropic and enthalpic contributions to the ECPs (resulting from the

reorganisation of the polymer matrix and the binding of the additive to the polymer) such that we can analyse the effect of of coarse-graining on these quantities.
This work is structured as follows: In sections 5.2 and 5.3 we will discuss the methods and the models used in this work. This is followed by a discussion of the computational details in section 5.4. In section 5.5, we will discuss the CG models for the different polymer-additive systems, which have been developed using the CRW method, and study how well these potentials are capable to predict ECPs. Furthermore, we will discuss the temperature transferability of the model and the structural properties of the polymer-additive systems. We will finally summarise the work in section 5.6.

5.2 Methodology

5.2.1 Conditional Reversible Work method

The non-bonded CG potentials used in this work are developed using the CRW method.[23] In the CRW method the interaction free energy between two groups of atoms is used as an effective pairwise CG interaction potential $U_{eff}(r)$. The calculation of the interaction free energy is performed using the thermodynamic cycle presented in Fig. 5.1. This figure shows the computation of the non-bonded interaction potential between a CG bead that represents the phenyl ring of a PS residue, and a CG bead that represents the phenyl ring in the EB molecule. $U_{eff}(r)$ denotes the effective CG pair potential, i.e the interaction free energy between the two groups of atoms when they are at the given distance r under the condition that they are embedded in their respective molecules. This free energy is calculated as the difference between the work of two different reversible processes. The first one ($RW(r)$) is the reversible work needed to bring the two groups of atoms (embedded in their molecule) from infinite distance to the distance r. The second one ($RW_{excl}(r)$) is the reversible work needed to perform the same process but neglecting the direct interaction between the two groups of atoms. The process labelled with $U_{eff}(\infty)$ in Fig. 5.1 denotes the free energy of the process of switching on the interaction between the two groups of atoms when they are at infinite distance. Since the two groups are not interacting at infinite distance the interaction free energy is zero: $U_{eff}(\infty) = 0$. According to the thermodynamic cycle, the effective CG pair potential can be calculated as $U_{eff}(r) = RW(r) - RW_{excl}(r)$.
Fig. 5.3a) shows the potentials corresponding to $RW(r)$, $RW_{excl}(r)$ and $U_{eff}(r)$ of

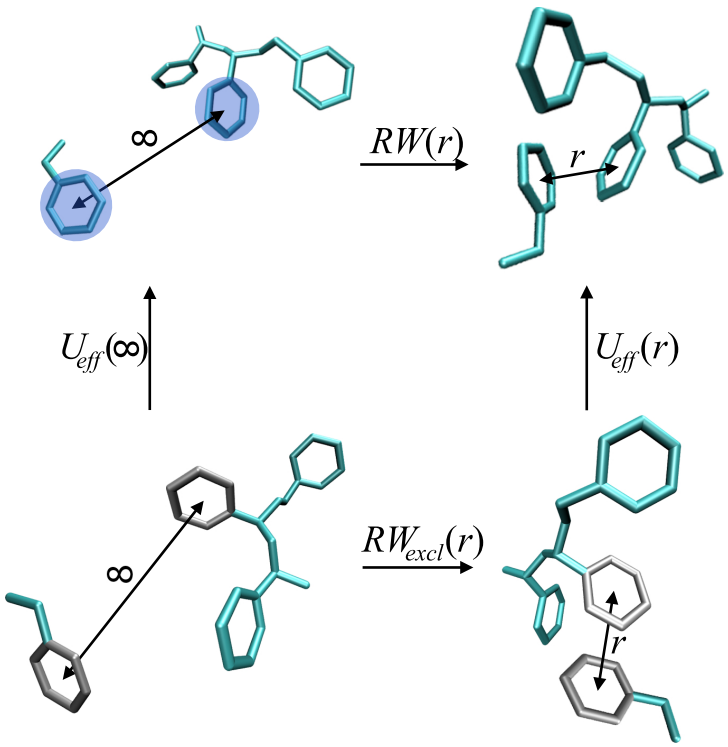


Figure 5.1: Thermodynamic cycle used to calculate the CRW CG potential for the interaction between the phenyl ring of PS and the phenyl ring of EB. The CRW pair potential ($U_{eff}(r)$) is defined as the conditional free energy associated with the process of introducing noncovalent interactions between the atoms in the two rings at a constrained distance r between their centers of mass. Conformational averaging at a selected temperature T is performed over all remaining degrees of freedom. Conditional free energy denotes the fact that conformational averaging is performed with the two atom groups embedded in their immediate chemical environment of the larger molecule. The CRW potential can be calculated as the difference between two reversible work potentials that include ($RW(r)$) and exclude ($RW_{excl}(r)$) the direct noncovalent ring-ring interactions.

the thermodynamic cycle for the non-bonded CG interactions between ME and the phenyl ring of PS. $RW(r)$ is the reversible work associated with the process of bringing two molecules together by pulling between the centers of mass of the two groups of atoms. This potential includes indirect contributions due to all the interactions between all the atoms in the two molecules. These indirect interactions are reflected by the irregular shape of the potential. When the same property is calculated excluding the direct interactions between the two group of atoms, we obtain a potential that contains only indirect interactions ($RW_{excl}(r)$). The difference between these two potentials results in an effective potential $U_{eff}(r)$ between the two group of atoms that does not contain indirect contributions.

The CRW CG potentials used in this work are developed in the gas phase. In order to apply these potentials to polymer-additive systems, they need to be transferable to the condensed phase. Furthermore, the temperature transferability is an important aspect that we will consider in this work. The CRW CG potentials are developed at a certain state point (in this work at 503 K) and it will be shown later, in section 5.5, how transferable they are over a wide temperature range.

5.2.2 Excess chemical potentials

In order to analyse the transferability of the CRW-based CG potentials we studied the excess chemical potentials of three different additives in polystyrene melts. We calculated excess chemical potentials of EB, ME and NP in atactic PS melts on an atomistic and CG level using fast-growth thermodynamic integration (FGTI) [24] for the larger solutes (EB and NP) and test particle insertion (TPI) [25] for ME. Unlike ME, the bulkier NP or EB additives cannot be inserted without severe particle overlaps in the system. In this context, FGTI has proven to be a useful method. [1,26] In FGTI, multiple TI runs are performed, where initially all the interactions between the polymer and the additive are switched off. These interactions are in the course of the simulation slowly turned on with a finite rate and the additive is coupled to the polymer matrix. The nonequilibrium coupling work W_{AB} can be related to the ECP $\Delta\mu_{\mathrm{ex}}$ using Jarzynski's nonequilibrium work theorem. [24]

$$\begin{aligned}\Delta\mu_{\mathrm{ex}} &= -k_B T \log\langle \mathrm{e}^{-\beta W_{\mathrm{AB}}}\rangle_{\mathrm{A}} \\ &= -k_B T \log \int_{\infty}^{-\infty} \mathrm{e}^{-\beta W_{\mathrm{AB}}} P(W_{\mathrm{AB}}) dW_{\mathrm{AB}}\end{aligned} \tag{5.1}$$

The angular brackets indicate an averaging over a canonical ensemble of the initial state A, $\beta = (k_B T)^{-1}$ with k_B the Boltzmann constant and T the temperature, and $P(W_{\mathrm{AB}})$ denotes the nonequilibrium work distribution. Detailed information about this method can be found in the work of Hess et al. [26,27] and Fritz et al. [1]

5.2.3 Thermodynamic analysis

In order to understand how the coarse-graining procedure affects the thermodynamics of the system, different contributions to the ECP are analysed. We consider a

system at constant pressure and temperature. Solute insertion (additive sorption) in a polymer melt can be considered as a two-step process. First, (i) a cavity of suitable size and shape for inserting the additive is created inside the polymer, followed by (ii) introducing the binding interactions between the additive and the cavity. The free energy ($\Delta\mu_{\mathrm{ex}}$) of the overall process can be decomposed to obtain an entropic contribution (ΔS_{ex}) and an enthalpic contribution (ΔH_{ex}). These two components can be further subdivided according to the two elementary steps above.[28,29] ΔH_{ex} is then written as

$$\Delta H_{\mathrm{ex}} = \Delta H_{\mathrm{R}} + \Delta H_{\mathrm{B}} \tag{5.2}$$

where ΔH_{R} is the reorganisation enthalpy associated with the loss of cohesive interactions in the cavity formation process, and ΔH_{B} is the binding enthalpy associated with the energy gained in the second step where the additive-polymer binding interactions are introduced. Hence, $\Delta H_{\mathrm{R}} > 0$ and $\Delta H_{\mathrm{B}} < 0$. Similarly, the excess entropy can be split in two contributions[28,29]

$$\Delta S_{\mathrm{ex}} = \frac{\Delta H_{\mathrm{R}}}{T} + \Delta S_{\mathrm{ap}} \tag{5.3}$$

In the literature, ΔS_{ap} has been referred to as the fluctuation entropy[30] or the solute-solvent entropy[31] because it can be related (through statistical mechanics formulas) to fluctuations of the solute-solvent interaction energy. Here, we use the subscript 'ap' where 'a' denotes the additive and 'p' the polymer melt. In Van der Waals systems, the major contribution to ΔS_{ap} arises from excluded volume interactions. ΔS_{ap} is the entropic cost of solute insertion; it is always negative and quantifies the loss of entropy associated with the reduced phase space that the polymer is able to sample due to the presence of an additive. Formally, $\exp[\Delta S_{\mathrm{ap}}/k_B]$ can be interpreted as the probability to observe an empty, transient cavity in the polymer melt. This cavity has the polymer repeat units in the equilibrium positions and orientations appropriate for accommodating all chemical moieties of the additive molecule. The contribution $\Delta H_{\mathrm{R}}/T$ to the excess entropy accounts for changes in all other interactions (the polymer-polymer interactions) not involving the solute. At constant pressure, $\Delta H_{\mathrm{R}}/T$ is positive[28] and partly compensates the negative contribution from ΔS_{ap}.

The excess enthalpy ΔH_{ex} and excess entropy ΔS_{ex} are temperature derivatives of

the free energy, i.e. $\Delta S_{\mathrm{ex}} = -\left(\partial \Delta\mu_{\mathrm{ex}}/\partial T\right)_p$ and $\Delta H_{\mathrm{ex}} = \left(\partial \beta \Delta\mu_{\mathrm{ex}}/\partial \beta\right)_p$, and can therefore be obtained from the temperature dependence of $\Delta\mu_{\mathrm{ex}}$ at constant pressure. ΔH_{B} is the sum of the non-bonded interaction energies between the polymeric matrix and the additive and is obtained directly from the simulations. ΔH_{R} and ΔS_{ap} can then be calculated by applying Eq. 5.2 and Eq. 5.3.

The CG model is temperature transferable if the predicted temperature dependence of $\Delta\mu_{\mathrm{ex}}$ agrees with the prediction obtained with the parent atomistic model, i.e. a transferable model reproduces the excess entropy (ΔS_{ex}). We point out that there is no a priori reason to expect that the CG model reproduces ΔH_{ex} and ΔS_{ex} in agreement with the UA model. CRW pair potentials are not energies but free energies. This means that a part of the entropy associated with a pairwise molecular interaction in the UA model description is contained in the effective pair potential of the CG model. This 'interaction entropy' is unfavorable (interactions bias the sampling of available phase space) and compensates part of the energetic attractions between two chemical groups, leading to effective pair potentials with shallow minima in comparison to the parent UA model. Therefore, the interactions in the CG model are effectively weaker (i.e. potential energy minima are less deep) than those in the UA model, while the entropy in the CG model is effectively larger. These implicit entropy contributions in the effective CG pair potentials may however cancel in the thermodynamic quantities ΔH_{ex} and ΔS_{ex} since they appear in the terms on the right hand sides of Eqs. 5.2 and 5.3 with opposite sign.

5.3 Models

5.3.1 United-atom model

All united-atom (UA) MD simulations for calculating of thermodynamic and structural properties were performed using the TraPPe UA FF[32]. This FF is targeted to reproduce vapour-liquid equilibria of molecular liquids. The development of the CG potentials has been done using the TraPPe FF, which has been applied successfully to polymer-additive systems in previous work[1] and it has been shown that it predicts values for the ECPs over a wide range of temperatures close to the experimental values. The CG polystyrene model of Fritz et al.[4] is based on the all-atom model of Müller-Plathe et al.[33] All the equilibrated atomistic starting configurations of the PS melt of the 24 chains of atactic 96mers are obtained by inverse mapping of equili-

brated CG melts at different temperatures. Details can be found in the work of Fritz et al. [1]

5.3.2 Coarse-grained model

We have used the CRW method developed by Brini et al. [23] to develop polymer-additive CG potentials. The mapping scheme is of crucial importance since different mapping schemes can lead to different results with respect to the transferability and representability of the CG model. The EB CG model, developed in this work, is composed of two beads: the first one represents the ethyl substituent (A_{EB}) an the second one represents the phenyl ring (B_{EB}) (see Fig. 5.2). The CG mapping points of these two beads are chosen as the center of mass of the group of atoms that they represent. NP and ME are represented as a single CG site, with the mapping point in its center of mass. The size of the atomistic and the CG methane is directly comparable, since we used a united-atom force field for the atomistic description. The coarse-grained PS model was previously developed following a CRW approach by Fritz et al. [4] The PS unit is represented by a two-bead model, one describes the phenyl ring and the other one represents the backbone. The model is capable of simulating polymer chains with different tacticity and is able to reproduce different structural and thermodynamic properties, for further informations we refer to the original paper of Fritz and et al. [4]

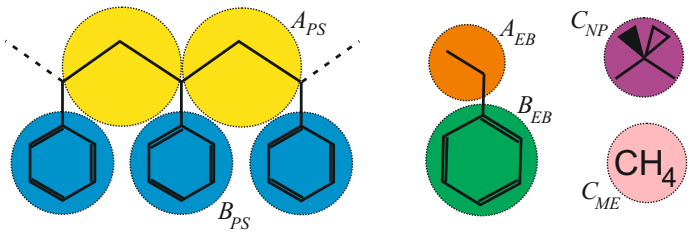


Figure 5.2: Mapping scheme for PS [4] and EB. NP and ME are both represented as a single bead in the CG representation.

Besides the non-bonded interactions, the bonded potentials are another crucial aspect of the CG model. ME and NP are mapped as single beads, therefore there is no bonded interaction needed. The CG EB is represented by two beads (A_{EB} and B_{EB}), this molecule is modelled as a rigid dumbbell, with the interatomic distance fixed at the equilibrium distance. For a description of the bonded interaction in the PS model, we refer the reader to the original work by Fritz and et al. [4]

5.4 Computational details

All simulations performed in this work were done using the GROMACS simulation package.[34]

5.4.1 CRW

Non-bonded CRW potentials are developed in vacuum through a series of distance constraint simulations between the centers of mass of the groups of atoms for which we want to calculate the CG interaction potential. The constraint algorithm employed is the linear constrain solver (LINCS).[35] The average constraint force is obtained from a 800 ns simulation trajectory generated employing a time step of 2 fs. The reversible work is then calculated integrating the average constraint force of every simulation over the constrain distance.[4] This distance is varied between 1.1 nm and 0.32 nm in steps of 0.02 nm. The constraint dynamics simulations are all performed at 503 K using a stochastic dynamics integrator with an inverse friction coefficient 0.5 ps. The cut-off distance for Lennard Jones interactions is 4.0 nm, this ensures that even at the largest sampling distance all the atoms of the two molecules are interacting. Since the sampling is carried out using a united atom force field no charges are involved.

As a CG mapping point for the additive the center of mass of the atoms that the CG beads represent is used. For the B_{PS} bead the center of mass of the central ring of a PS trimer is used. The interaction potentials of the additive with the A_{PS} bead is less straight forward to calculate. As it can be seen in Fig. 5.2 this CG bead represent a CH group and two $1/2$ CH groups. The interaction free energy for half atoms cannot be calculated. Therefore, we choose to define $RW_{excl}(r)$ as the potential that results from the arithmetic average between a $RW_{excl}(r)$ calculated in a process where only the direct interaction of the additive with the central CH is excluded and another $RW_{excl}(r)$ that characterises a process where all the interaction between the additive and the atom of A_{PS} are excluded. We note that this can cause uncertainties in the enthalpic interactions. To obtain the second series of interaction potentials for the NP-PS interaction, a pentamer of PS has been employed instead of a trimer.

5.4.2 Excess chemical potentials

All simulations are performed under NpT conditions at a pressure of 1 atm. It has been shown that under these condition the Fritz PS CG model is able to reproduce the thermal expansion coefficient, although the densities of the CG and UA model are slightly different[1,4]. The atomistic PS systems used in this work for the calculation of ECPs are obtained by inverse mapping of equilibrated coarse-grained melts. The calculations have been done in a temperature range between 503 K (polymer melt) and 383 K (close to the experimental T_g of 373 K). The T_g of the CG PS model is 363 K.[36] For the pure polystyrene matrix we performed simulations of 16 ns at each temperature. The starting configurations for the FGTI calculations were taken from the last 8 ns. The ECPs of NP and EB in the atomistic and CG systems are calculated by performing 50 independent FGTI calculations. The overall coupling time of the additive is 2 ns, in which the Lennard Jones interactions are switched on. For the FGTI runs we used a Langevin thermostat with a friction coefficient of 1 ps^{-1}, to obtain a canonical distribution when the additive is nearly decoupled from the PS matrix. The ECPs of ME in PS of the atomistic and CG systems were calculated using TPI. 3000000 insertions are performed every 10000 timesteps in an overall 20 ns PS trajectory, obtained at different temperatures using a Berendsen[37] thermostat with a coupling time of 1ps.

5.4.3 MD simulation

We have performed MD simulations at 503 K in order to calculate radial distribution functions (RDFs) for the three different polymer-additive systems. The systems contain a melt of 24 chains (96mers) of atactic PS and 10 additive molecules. The simulations are performed under isothermal-isobaric conditions using a Parrinello-Rahman barostat[38,39] with a coupling time of 1 ps and a Nose-Hoover thermostat[40,41] with a coupling time of 0.1 ps. The bonds are constrained using the LINCS algorithm.[35]. The UA simulations are 60 ns long using an integration time of 2 fs. The CG simulations are 40 ns long using the same integration time step.

5.5 Results and discussion

We studied the transferability of CRW CG potentials for different polymer-additive systems. To this end, structural and thermodynamic properties are calculated for

the following three systems: ethylbenzene, methane and neopentane in polystyrene. In this work, we have studied the ECPs of the additives in a melt of 24 chains of atactic 96mers of PS. Firstly, we will present the CRW CG potentials developed in this work, then we discuss the temperature dependence of the ECPs for the three different systems to then, in a later section, discuss the structure of the system and analyse the different enthalpic and entropic contributions to the ECPs.

5.5.1 CG potentials

Using the CRW method we obtained CG non-bonded interaction potentials between the additives and the polymer. These potentials are reported in Fig. 5.3 in panels b), c) and d). All the potentials have a monotonically varying tail, which is a clear indication that the potentials are not containing any multi-body contributions[23]. Also the relative magnitudes of the interaction potentials between the additives and A_{PS} and B_{PS} are reasonable. In fact, all the interaction potentials in which B_{PS} is involved are deeper and show a bigger excluded volume compared to the one where A_{PS} is involved.

Fig. 5.3c) shows the interaction between EB and PS. The interaction $B_{EB} - B_{PS}$ is the deepest of the four; this reflects the bigger size and the stronger interaction that phenyl rings have compared to few atoms of the backbone. It is also interesting to note that the potential $A_{EB} - B_{PS}$ and $B_{EB} - A_{PS}$ (dashed lines) are to some extent similar, since they both represent the interaction between alkyl chains with a phenyl ring.

Fig. 5.3b) shows two sets of potentials for the interaction between ME and PS. The difference between these two sets is the distance at which the interaction between the beads is considered negligible, for the first set (continuous lines) this distance is 1.1 nm and for the second set (dashed lines) it is 1.0 nm. The difference between the two sets is minimal but its effect on the ECP is not negligible, as we will see later in section 5.5.2). This gives us an idea of how small variations in the potential (due to any possible cause, even statistical errors) can influence the computed ECPs.

In Fig. 5.3d) the interaction potentials between NP and PS are reported. These are shifted to bigger distance and are deeper compared to the interaction potential between ME and PS, reflecting the bigger size of NP molecule. In order to pinpoint eventual sampling issues during the CRW procedure in vacuum, two sets of potentials have been calculated. The first set (continuous line) refers to a model parametrised using a 3mer of PS, the second set (dashed line) refers to a model parametrised using

a 5mer of PS. The interaction potentials are similar, but the potentials obtained with the 5mer are smoother.

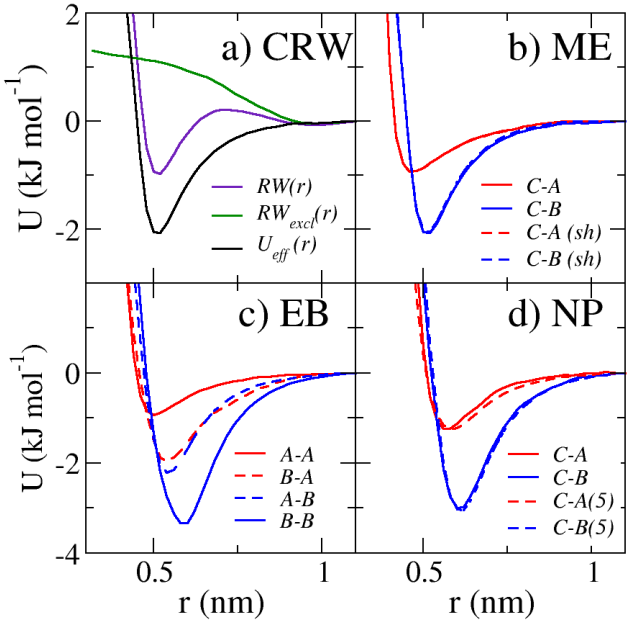


Figure 5.3: Non-bonded interaction potentials developed in this work. Panel a) describes how the CRW method works taking as example the interaction $C_{ME} - B_{PS}$. Panels b), c) and d) show the interaction potentials between PS and ME, EB and NP, respectively. In the legends of these three panels the first letter denotes the additive bead and the second letter denotes the PS bead to which the interaction is referred; also all the interactions with A_{PS} are coloured in red and with B_{PS} are coloured in blue. In b) and d) dashed lines refers to a second set of potentials developed respectively considering a cutoff of $1.0nm$ and using during the development of interaction potential a pentamer of PS instead of a trimer.

5.5.2 ECPs of additives in polystyrene melts

Fig. 5.4 shows the ECPs for the three different polymer-additive systems. The black dots show the UA ECPs and the red triangles are the results from the CG simulations. The dashed lines show the temperature dependence of the ECPs. The slopes of these linear fits provide the excess entropies. As shown in previous work, the UA force field for EB is capable of predicting ECPs close to the experimental values. [1] Quantitative differences are observed between the ECPs obtained with the CG and reference atomistic models. The differences will be discussed below. Interestingly, the temperature dependence of the ECPs is the same for the CG and UA models, indicating that the

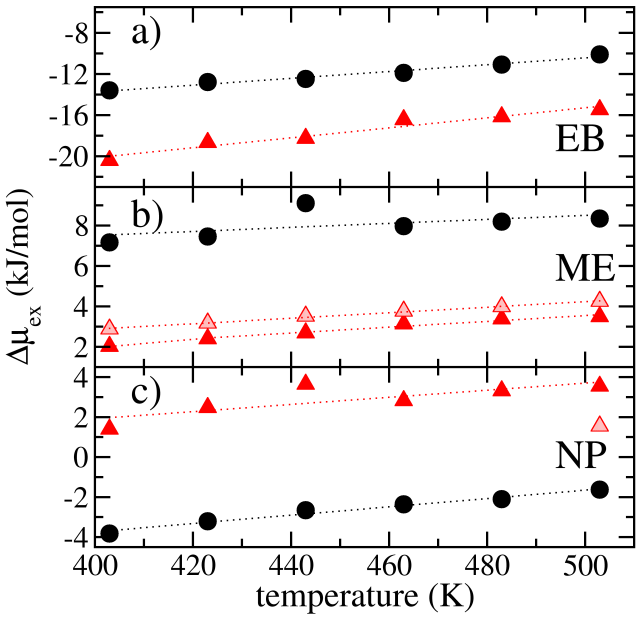


Figure 5.4: Temperature dependence of the excess chemical potentials of EB, ME and NP in a PS melt. The black dots show the UA simulation results, the red triangles the CG simulation results. Linear regressions are shown as dashed lines. The red patterned triangles in panel b) show results obtained with the potentials C-A(sh) and C-B(sh) in Fig. 5.3b). The red patterned triangle in panel c) shows the result obtained with the potentials C-A(5) and C-B(5) in Fig. 5.3d).

CG model is temperature transferable and the excess entropies of these additive are correctly represented by the CG model. Fig. 5.4a) shows that the CG simulations predict too low values for the ECPs of EB in PS. This also happens in the case of ME (shown in Fig. 5.4b)). The ECPs of ME are positive, indicating low solubility of ME at these temperatures. Contrary to what is observed in a) and b), Fig. 5.4c) shows that the UA simulations predict negative values for the ECPs of NP and larger, positive values at the CG level. We will discuss this inverse trend in section 5.5.4, where we discuss in greater detail enthalpic and entropic contributions to the ECPs.

Shen et al. have shown before that small changes in the CG potentials can lead to big changes in the thermodynamic properties.[21] We address this issue by varying the CG potentials used in this work and study the effect on the ECPs. As described in section 5.5.1, we have derived two CG potentials for the interaction between PS and ME using two different cutoffs. Fig. 5.4b) shows the results for the ECPs derived with a cutoff of 1.1 nm (red filled triangles) and with a cutoff of 1.0 nm (red patterned

triangles). The potentials are reported in Fig. 5.3b). The difference between the potentials is almost not visible. However, small changes in the CG potentials can lead to relatively large differences in the ECPs. The differences in the potentials are of the order of magnitude of the errors that occurs when deriving the potentials. The ECPs for the CG potential with a cutoff of 1.0 nm are shifted around 1 kJ/mol in comparison to the one calculated using the CG potential with a cutoff of 1.1 nm. In principle, this opens a route to tailor the CG potentials such that they are in perfect agreement with the UA values. Since already small changes in the potentials are enough to achieve this, it will not have any significant effect on the structural properties. As previously mentioned in section 5.5.1, we have also developed a second set of potentials for the interaction of NP with PS. This second set is derived using a 5mer of PS for the vacuum sampling instead of a trimer. In general, the vacuum sampling of the polymer strand of a given length is only an approximation to describe the interaction with the polymer. The isolated PS strand is normally embedded in a polymer chain and could therefore sample slightly different conformations than when being part of the longer polymer chain. This might lead to structural or thermodynamic discrepancies between the UA reference system and the CG model. Fig. 5.3d) shows the different CG potentials. The CG potentials that are based on the 5mer sampling show only a small shift to larger distances. The corresponding ECP shown in Fig. 5.4c) (red patterned triangle) however differs significantly from the value obtained with the 3mer potential. The ECP using the CG potential obtained with the 5mer sampling is about 2 kJ/mol lower and therefore closer to the UA reference. This shows that on the one hand the ECP is very sensitive to small changes in the interaction potential. On the other hand, it indicates that the way the CG potentials are derived can strongly influence the transferability and representability of the model. Under these conditions, it is even more striking that such simplified CG models can predict sensitive thermodynamic quantities, like the ECPs, in the right order of magnitude and with the correct temperature dependence.

5.5.3 Structure

In order to compare the ECPs of the UA and CG system it is mandatory that the local chemical environments of the additives in the UA and CG polymer matrix are similiar. For this reason we calculated RDFs for the three different systems at the UA and CG level at 503 K. Fig. 5.5 and 5.6 show the radial distribution functions of EB with PS, and ME and NP with PS, respectively. The UA and CG RDFs are in a good agreement,

considering that this property is a prediction of the CG model that we remind is not parametrized to reproduce any properties. However, the CG RDFs are slightly shifted to smaller distances. This is probably due to the spherical representation of slighly anisotropic groups. Because the ring structure of the phenyl group of PS is modeled as a sphere, the interaction potentials between B_{PS} and the additives represent averages of additive-ring face and additive-ring side interactions. This causes the effective excluded volume of the beads to be an average between these two limiting configurations, and this can lead to a small shifting in the calculated RDF. We note that the shifting in the RDFs involving B_{EB} in Fig. 5.5 b) and d) is more pronounced compared to the shifting of the other bead pairs reported in the same figure and in Fig. 5.6.

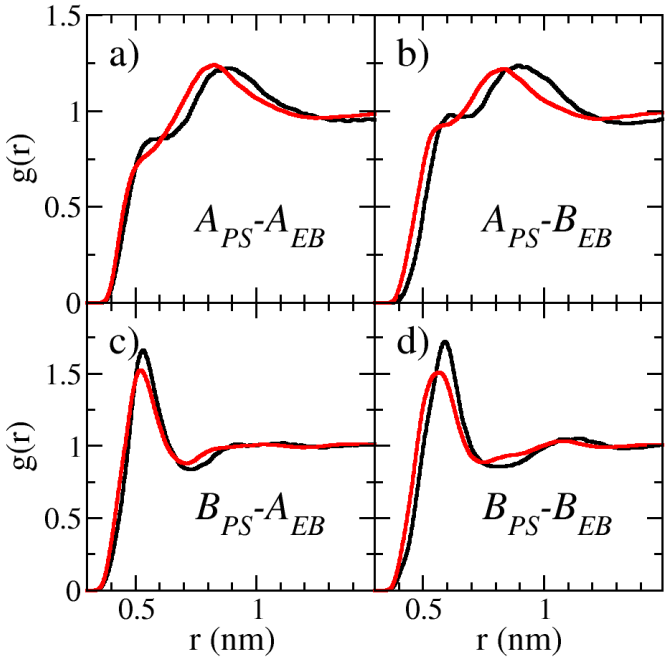


Figure 5.5: Radial distribution functions of the PS CG beads A_{PS} and B_{PS} with the EB CG beads A_{EB} and B_{EB}. The black lines are obtained from the UA simulations and the red lines form the CG simulations using the CRW-derived CG potentials.

The comparison of the UA RDFs and CG RDFs indicate that the chemical environment of the additives is similar in both systems. Therefore, the UA and CG ECPs can be meaningfully compared and deviations can be related to details of the CRW potentials.

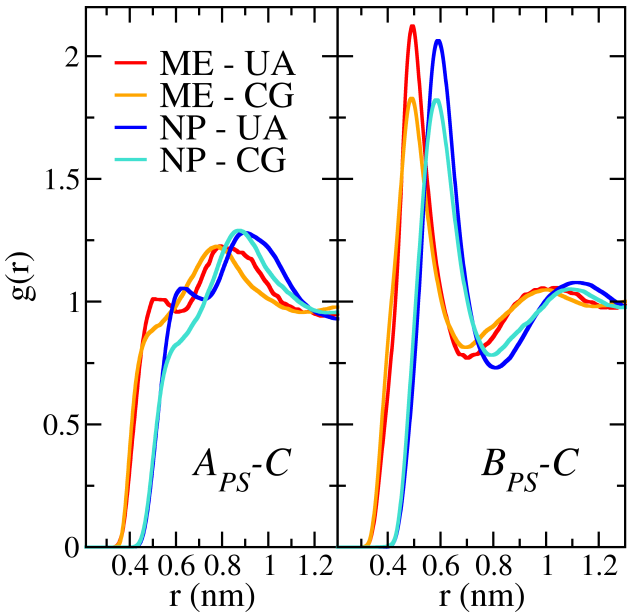


Figure 5.6: Radial distribution functions of the PS CG beads A_{PS} and B_{PS} with the coarse-grained ME, beadtype C_{ME}, and with the coarse-grained NP, bead type C_{NP}. The red and blue lines are obtained from the UA simulations and the orange and turquoise lines form the CG simulations using the CRW derived CG potentials.

5.5.4 Entropy and enthalpy contributions to ECPs

In order to analyse discrepancies in the ECPs in greater detail, we studied the excess enthalpies and excess entropies and their thermodynamic contributions to the UA and the CG systems, as described in section 5.2.3. All data are reported in Tab. 5.1, where the arrows illustrate the direction of the trend between UA and CG quantities. All values are reported at the temperature of 503 K in units of kJ/mol. Firstly, we discuss the excess enthalpy (ΔH_{ex}) and the excess entropy (ΔS_{ex}) of the three different systems and after that we analyse the different contributions to them. Fig. 5.4 and the data in Tab. 5.1 show that a reasonable agreement is achieved in the temperature dependence of the ECPs with an approximately constant offset between the UA and CG data. Hence, the excess entropies ΔS_{ex} of the additives are reproduced with the CG model while the mismatch between the UA and CG ECPs is of enthalpic origin (ΔH_{ex}). The three different additives show different trends (indicated by the arrows in Tab. 5.1). In the case of NP, the CG model predicts systematically too high ECPs (where in the case of EB and ME the ECPs are too low).

A comparison of the UA and CG predictions of ΔH_{ex} and $-T\Delta S_{ex}$ in Table 1 shows that the largest, absolute discrepancies are obtained with EB. In order to better understand these discrepancies we have decomposed ΔH_{ex} and $-T\Delta S_{ex}$ according to the scheme discussed in section 5.2.3. If we compare the relative changes in binding enthalpies among ME, NP and EB (the order of increased binding interaction) at both the UA and CG level, we find that the CG model is consistent with the UA model. For example, the ΔH_B associated with the insertion of a EB molecule in the polymeric matrix is about 4 times the ΔH_B associated with the insertion of a ME molecule, both in the UA and in the CG simulations. This observation is very interesting since it indeed indicates that the newly developed CRW potentials provide a chemically realistic description of changes in ΔH_B upon varying chemical groups in condensed phase systems. A similar comparison made for the relative changes in reorganization enthalpies among ME, NP and EB (the order of increasing reorganization enthalpy as predicted by the UA model) however clearly points out that ΔH_R obtained for EB with the CG model is unrealistically small. As a result, ΔH_{ex} of EB, predicted by the CG model, is underestimated (too negative) in comparison with the UA model. This observation is consistent with the structural analysis in section 5.3 which indicated that the spherical description of the phenyl bead yields a CRW potential with too small excluded volume repulsion. The excess entropy contains two contributions (Eq.5.3), which, at constant pressure, have opposite sign. If ΔH_R is predicted too small, we expect that ΔS_{ap} is also predicted too small; i.e. if the enthalpy cost of cavity formation (ΔH_R) becomes smaller, the probability increases that thermal fluctuations lead to the formation of transient cavities. This can be observed in the last column of Table 1 where, in comparison to the data of the UA model, $-T\Delta S_{ap}$ of EB(CG) is not significantly bigger than that of NP(CG). This leads to an underprediction of the excess entropy for EB.

We finally note that in the CG system ΔH_R, ΔH_B and ΔS_{ap} are strongly dependent on the mapping scheme and therefore on the number of degrees of freedom that are lost in the CG procedure. It therefore is difficult to quantitatively compare these quantities for the three different additives, which all map a different number of atoms in an effective CG interaction site. Although the contributions of ΔH_R, ΔH_B and ΔS_{ap} to the excess entropies and enthalpies add up to reasonably accurate predictions in some cases (e.g. the excess entropy of ME and NP), it remains unclear why the excess enthalpy of NP is overestimated, while being underestimated for ME and EB. This may be related to geometric aspects of the mapping scheme that we em-

ployed. However, to address these questions, a systematic study of a single system with different CG mapping schemes is required where enthalpy-entropy compensation of implicit entropy contributions in the effective interaction potentials is studied in greater detail (see discussion of Eqs. 5.2 and 5.3 in section 5.2.3).

	$\Delta\mu_{ex}$	ΔH_{ex}	$-T\Delta S_{ex}$	ΔH_R	ΔH_B	$-T\Delta S_{ap}$
EB (UA)	-10.1 ↑	-26.8 ↑	16.6 ↓	35.5 ↑	-62.3 ↓	52.2 ↑
EB (CG)	-15.5	-40.0	24.7	10.2	-50.2	34.7
ME (UA)	8.3 ↑	3.0 ↑	5.0 ↓	18.3 ↑	-15.3 ↓	23.6 ↑
ME (CG)	3.5	-4.2	7.7	8.3	-12.5	16.0
NP (UA)	-1.6 ↓	-12.2 ↓	10.6 ↑	30.7 ↑	-42.9 ↓	41.3 ↑
NP (CG)	3.5	-5.8	9.1	21.1	-26.9	30.4

Table 5.1: Excess chemical potentials ($\Delta\mu_{ex}$), excess entropies (ΔS_{ex}), excess enthalpies (ΔH_{ex}), reorganisation enthalpy (ΔH_R), binding enthalpy (ΔH_B) and the solute-solvent entropy (ΔS_{ap}) for three systems studied in this work: EB, ME and NP in PS. We report the values for the UA and the CG system at 503 K. The arrows illustrate the direction of the trends between the CG and UA quantities. All units are in kJ/mol.

5.6 Conclusions

We have studied the thermodynamic transferability of CRW derived CG potentials for three different polymer-additive systems. We have shown in this work that these CG potentials are capable of predicting the correct structural correlations in the mixture. Furthermore, the ECPs of the additives obtained with the UA and the CG models are in acceptable agreement, given that this quantity is very sensitive to small changes in the potentials. It is striking that the CRW-based CG models show a good temperature transferability. This means that the excess entropies are well reproduced, while discrepancies are observed in the excess enthalpies.

Although we already achieve a relatively good agreement with the UA reference system, one can try to improve the models even further. Theoretically, CRW-based CG models can be tailored such that thermodynamic quantities, like the ECPs, are in excellent agreement with the atomistic reference system. We have shown that a small shift in potential can cause a relatively big shift in the ECPs. Since this shift is very small, structural properties will not be affected. This opens up a way to simulate chemistry-specific polymer-solvent systems with coarse-grained models on large length and time scales, examples of which include plasticizers in polymer networks, polymer swelling and dissolution in specific solvents, polymer wetting processes, etc.

5.7 Bibliography

[1] D. Fritz, C. R. Herbers, K. Kremer, and N. F. A. van der Vegt. *Soft Matter*, 5:4556, 2009.

[2] S. Pandiyan, D. Brown, S. Neyertz, and N. F. A. van der Vegt. *Macromolecules*, 43:2605, 2010.

[3] V. A. Harmandaris, N. P. Adhikari, N. F. A. van der Vegt, and K. Kremer. *Macromolecules*, 39:6708, 2006.

[4] D. Fritz, V. Harmandaris, K. Kremer, and N. F. A. van der Vegt. *Macromolecules*, 42:7579, 2009.

[5] D. Fritz, K. Koschke, V. A. Harmandaris, N. F. A. van der Vegt, and K. Kremer. *Phys. Chem. Chem. Phys.*, 13:10412, 2011.

[6] H. Eslami, H. A. Karimi-Varzaneh, and F. Müller-Plathe. *Macromolecues*, 44:3117, 2011.

[7] L. Delle Site, C. Holm, and N. F. A. van der Vegt. *Top. Curr. Chem.*, 307:251, 2012.

[8] C. Li, J. Shen, C. Peter, and N. F. A. van der Vegt. *Macromolecules*, 45:2551, 2012.

[9] Florian Müller-Plathe. *Chem. Phys. Chem.*, 3:754, 2002.

[10] C. Chen, P. Depa, J. K. Maranas, and V. G. Sakai. *J. Chem. Phys.*, 128:124906, 2008.

[11] P. Carbone, H. A. Karimi-Varzaneh, X. Chen, and F. Muller-Plathe. *J. Chem. Phys.*, 128:064904, 2008.

[12] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *J Phys. Chem. B*, 111:7812, 2007.

[13] K. A. Maerzke and J. I. Siepmann. *J. Phys. Chem. B*, 115:3452, 2011.

[14] J. D. McCoy and J. G. Curro. *Macromolecules*, 31:9362, 1998.

[15] B. M. Mognetti, P. Virnau, L. Yelash, W. Paul, K. Binder, M. Müller, and L. G. MacDowell. *J. Chem. Phys.*, 130:44101, 2009.

[16] D. Reith, M. Pütz, and F. Müller-Plathe. *J. Comput. Chem.*, 24:1624, 2003.

[17] A. Lyubartsev and A. Laaksonen. *Phys. Rev. E*, 52:3730, 1995.

[18] F. Ercolessi and J. Adams. *Europhys. Lett.*, 26:583, 1994.

[19] N. Zacharopoulos, N. Vergadou, and D. N. Theodorou. *J. Chem. Phys.*, 122:244111, 2005.

[20] A. Villa, N. F. A. van der Vegt, and C. Peter. *Phys. Chem. Chem. Phys.*, 11:2068, 2009.

[21] J. W. Shen, L. Li, N. F. A. van der Vegt, and C. Peter. *J. Chem. Theory Comput.*, 7:1916, 2011.

[22] Y. Wang, W. G Noid, P. Liu, and G. A. Voth. *Phys. Chem. Chem. Phys.*, 11:2002, 2009.

[23] E. Brini, V. Marcon, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 13:10468, 2011.

[24] C. Jarzynski. *Phys. Rev. Lett.*, 78:2690, 1997.

[25] B. Widom. *J. Chem. Phys.*, 39:2802, 1963.

[26] B. Hess, C. Peter, T. Ozal, and N. F. A. van der Vegt. *Macromolecules*, 41:2283, 2008.

[27] B. Hess and N. F. A. van der Vegt. *Macromolecules*, 41:7281, 2008.

[28] C. Peter and N. F. A. van der Vegt. *J. Phys. Chem. B*, 111:7836, 2007.

[29] N. F. A. van der Vegt, V. A. Kusuma, and B. D. Freeman. *Macromolecules*, 43:1473, 2010.

[30] D. Ben-Amotz, F. O. Raineri, and G. Stell. *J. Phys. Chem. B*, 109:6866, 2005.

[31] N. F. A. van der Vegt and W. F. van Gunsteren. *J. Phys. Chem. B*, 108:1056, 2004.

[32] C. D. Wick, M. G. Martin, and J. I. Siepmann. *J. Phys. Chem. B*, 104:8008, 2000.

[33] F. Müller-Plathe. *Macromolecules*, 29:4782, 1996.

[34] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Theory Comput.*, 4:435, 2008.

[35] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. *J. Comput. Chem.*, 18:1463, 1997.

[36] V. Marcon, D. Fritz, and N. F. A. van der Vegt. *Soft Matter*, 8:5585, 2012.

[37] H. J. C. Berendsen, J. P. M Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak. *J. Chem. Phys.*, 81:3684–3690, 1984.

[38] M. Parrinello and A. Rahman. *J. Appl. Phys.*, 52:7182, 1981.

[39] S. Nose and M. L. Klein. *Mol. Phys.*, 50:1055–1076, 1983.

[40] S. Nose. *Mol. Phys.*, 52:255, 1984.

[41] W. G. Hoover. *Phys. Rev. A*, 31:1695, 1985.

5.8 Supplementary Informations

5.8.1 Half atom groups and free energy calculations

As we pointed out in section 5.4.1 the interaction free energy between A_{PS} and the other beads can not be calculated since this bead of PS is representative of one full CH_2 group and two half CH groups. We therefore defined the interaction between such bead an the other ones as the average between two potentials calculated including or neglecting the interactions of the two CH groups with the additives. In order to understand how much this choice influences the ECPs we compare the calculation of the ECP of NP at 503K employing three different sets of interaction potential $A_{PS}-C_{NP}$. The first one (A_{PS}^{off}) is developed considering only the interaction between NP and the backbone CH_2, in the second one (A_{PS}^{on}) is developed considering the full interaction between NP and the backbone $CH-CH2-CH$, and the third one (A_{PS}) is the arithmetic average of the two and it is the one employed for all the other calculations of this paper. For the tree sets the mapping point is the center of mass of the atoms composing the A_{PS} bead, that is calculated considering the full mas of the CH_2 group and half of the mass of the two CH groups. A comparison between the potential is shown in Fig. 5.7.

As it is possible to see the three potentials present different depth of the minima; the minima become deeper as the number of atoms considered part of the A bead increases. It is interesting to note that, since the interaction potentials is defined as a *conditional* free energy, the excluded volume of the three models is the same. This is due to the fact that the three models have the same mapping point and the NP can only approach the PS backbone almost perpendicularly and toward the CH_2 group, since other directions are sterically hindered due to the presence of the backbone and of the phenyl rings.

In Tab. 5.2 are compared The ECPs calculated with the three models at 503K. Considering the fact that the excluded volumes for the interaction $A_{EB}-C_{NP}$ of the three models are similar we can argue that the biggest contribution to the differences between the ECP obtained with the different models is related with the binding enthalpy associated with the additive-backbone interaction (ΔH_B). According to this A_{PS}^{off} and A_{PS}^{on} set the limits in between which sits the calculated ECP of any model built assuming as interaction potential a combination of these two. In this case since A_{PS}^{off} is the less deep of the three interaction potentials its enthalpic contribution to

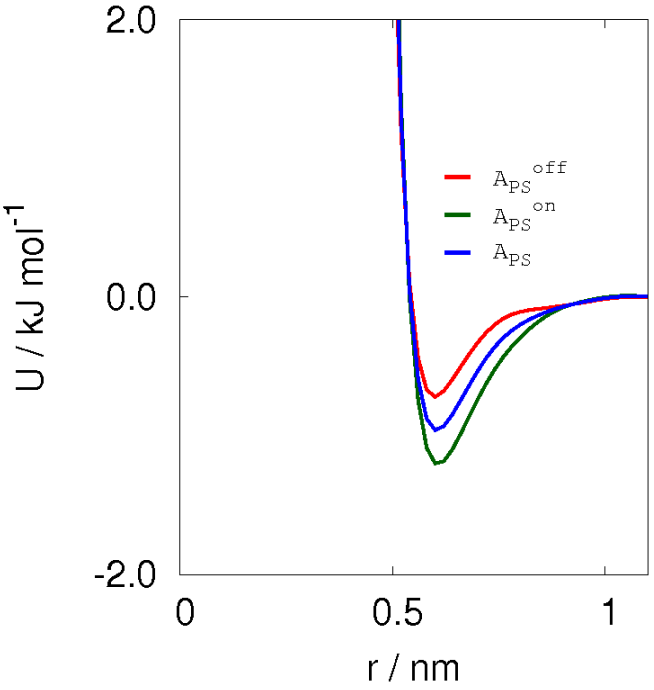


Figure 5.7: Comparison between three interaction potentials representing the interaction between the backbone bead of PS and NP. A_{PS}^{off} considers only the contribution to the potential of the central CH_2 group of the backbone repeating unit, A_{ps}^{on} considers the full contribution of the backbone repeating unit to the potentials, and A_{PS} is the arithmetic average between the two and it is the one employed for all the other calculation in the paper.

	$\Delta\mu_{\text{ex}}$
A_{PS}^{off}	+5.4
A_{PS}	+3.5
A_{PS}^{on}	+1.5

Table 5.2: Excess chemical potentials ($\Delta\mu_{\text{ex}}$) calculated for inserting a molecule of NP in PS matrix at 503K. The three different models employed consider different contribution of the two CH groups to the interaction potential $A_{PS} - C_{NP}$. In particular A_{PS}^{off} neglect this contribution, A_{PS}^{on} consider their full contribution and A_{PS} consider a partial contribution, since this interaction potential is defined as the arithmetic average of the other two. The ECPs value are in unit of $kJmol^{-1}$

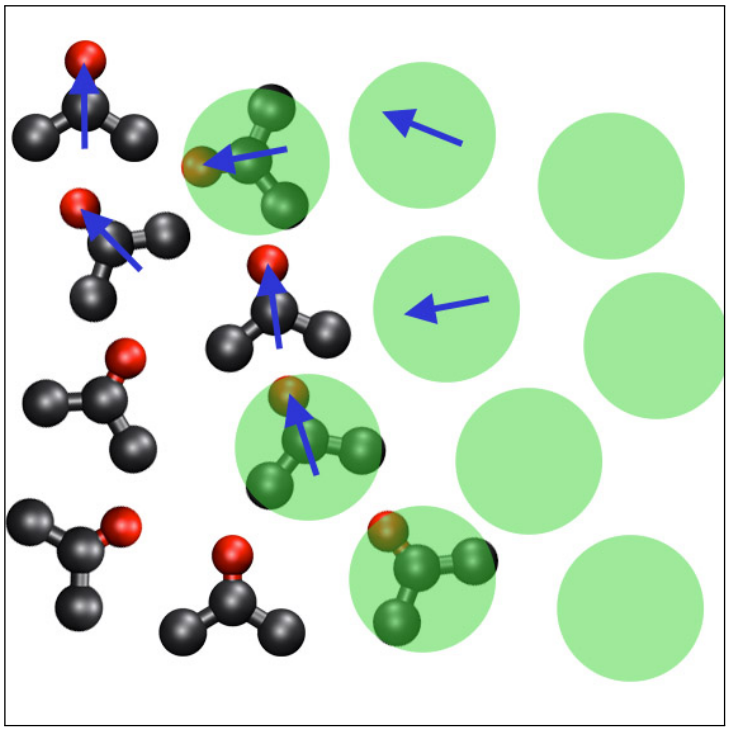
the ECP (ΔH_B) is the smaller one in absolute value, leading to a bigger value of the ECP. On the contrary A_{PS}^{on} that is the deeper interaction potentials expresses the smaller ECP. It is interesting to note that since A_{PS} in our case is assumed to be the arithmetic mean of the two potentials its ECP sits in the middle of between the two limiting ECPs. This supports our hypothesis about the contribution of these poten-

tials mainly to the enthalpic part of the ECP.

It has to be noted that the choice of the interaction potential to employ is critical, since in principle it can lead to differences between the calculated ECPs of 4 $kJmol^{-1}$. Nevertheless it is reasonable to assume that if the interaction free energy between A_{PS} and C_{NP} could be calculated it will be similar to the average of the two limit configurations. Therefore we speculate that the systematic error in the calculation of the ECPs associated with this issue is smaller than the $4kJmol^{-1}$ difference between the ECPs of the two limit cases.

6 Limits in the applicability of Conditional Reversible Work pair potential

In this work we investigate the possibility of employing coarse grained pair potentials developed with the conditional reversible work method to describe interactions between polar molecules. This class of compounds is characterized by directional multibody interactions and it is not a priori clear if the use of pair potentials will lead to models able to reproduce any property. We try to understand the limit of this approach by studying the behavior of single site CG models for a series of molecular liquid with increasing polarity. In particular we discuss limitations linked to the nature of the interactions and also technical limitations due to the method used to calculate conditional reversible work potentials.



6.1 Introduction

Simplifying the representation of molecules by lumping toghether groups of atoms in *beads* enables the possibility of studing phenomena happening at a time and size scale non accessible to standard atomistic molecular dynamics (MD). Two different apporaches are available in order to obtain simplified models still able to accurately describe the chemistry of the systems. The first one is a top down kind of approach that paramterizes the interaction between beads in order to reproduce macroscopic thermodynamic properties in a similar whay to what is done with many atomistic force field.[1–3] The second one is a bottom up approach that uses informations from a more detailed level of simualtion (e.g. atomistic level) in order to develop CG interaction potential.[4–12] In between this second class we recently proposed the Conditional Reversible Work (CRW) method,[12] that defines the interaction potential between two beads as the interaction free energy between the two groups of atoms represented by the beads calculated under the condition that they are embedded in their respective molecules. This methods delivers pair potentials free of indirect contributions in contrast to many other CG approaches[4–9] that deliver *effective* pair potentials that include multibody informations. In principle *effective* pair potentials enable the possibility to simulate systems where multibody interactions are important employing a computationally cheap pair description of the interactions, but the inclusion of state point dependent informations into the potentials limits the ability of the model to predict properties of the systems at a different state point from the one used during its development. It is therefore interesting to explore the limit of the applicability of pair potentials to systems characterized by multibody interactions. In this work we compare the property of parent atomistic models and single bead CG models of a series of fluid of increasing polarity (toluene, dimethyl ether, acetone, dimethyl sulfoxide), in order to understand when the approximation of using pair interactions becomes too severe to be employed.

6.2 Conditional Reversible Work Method

6.2.1 Sampling of the interaction in bulk liquid

The non-bonded CG potentials used in this work are developed using the CRW method.[12] In this method the interaction free energy between groups of atoms is used as an effective pair CG interaction potential $U_{eff}(r)$.

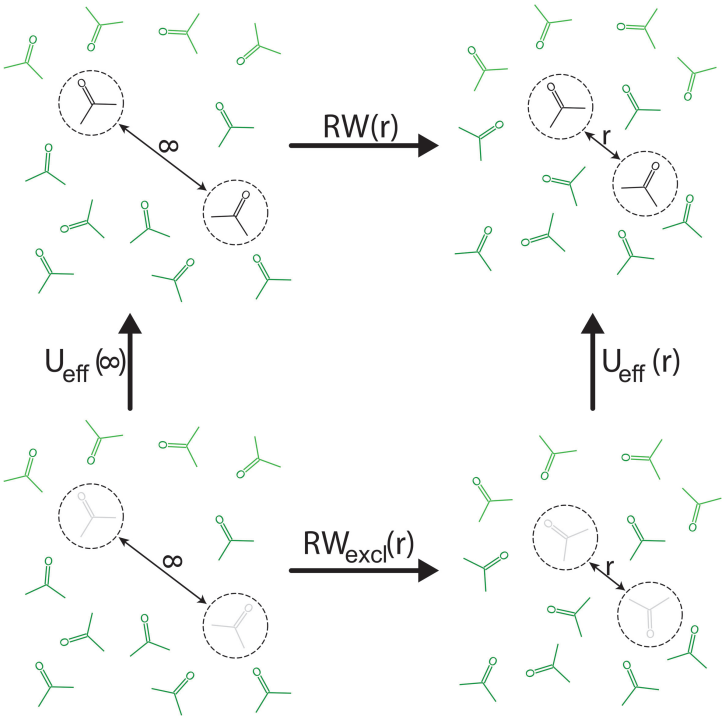


Figure 6.1: Thermodynamic cycle used to calculate the CRW CG potential between two beads representing two molecules of acetone. In this example the sampling is carried in bulk acetone, in this way the configurational space explored by the two molecules is the one that characterizes acetone liquid state. The CRW interaction potential ($U_{eff}(r)$) is the conditional interaction free energy associated with the process of introducing non-covalent interactions between the atoms of the two acetone molecules when they are at the distance r. $U_{eff}(r)$ can be calculated as the difference between the reversible works associated with two processes where the two molecules are reversibly pulled together from infinite distance to the distance r. The difference between the two is that the first process ($RW(r)$) includes the direct interaction between the two molecules, while the second one ($RW_{excl}(r)$) does not account for them.

The calculation of the interaction free energy is performed using the thermodynamic cycle presented in Fig. 6.1. There it is shown the computation of the non-bonded interaction potential for two CG beads representing two acetone molecules.

In this case the sampling is carried in bulk acetone, in this way the relative orientations of the molecules that contribute to the interaction potential are the ones characterizing the phase space of bulk acetone. Since the free energy associated with switching on the direct interaction between the two molecules when they are at infinite distance ($U_{eff}(\infty)$) is 0, the interaction free energy $U_{eff}(r)$ can be calculated as the difference $RW(r) - RW_{excl}(r)$. $RW(r)$ is the reversible work associated with the process of pulling together the two acetone molecules from infinite distance to the distance r. $RW_{excl}(r)$ is the reversible work associated with a process similar to the previous one, but with the difference that this time the direct interactions between the two molecules are neglected. This thermodynamic cycle shows also that $U_{eff}(r)$ is a pair potential free of indirect contributions, in fact $RW(r)$ contains both direct and indirect contributions to the potential, while $RW_{excl}(r)$ contains only the indirect ones. It is clear than that from the difference of these two is obtained an interaction potential that represents only the direct pair interaction between the two molecules.

6.3 Models

6.3.1 Atomistic models

In this work we studied the behavior of single bead CG models of four different molecular liquids with increasing polarity: Toluene (TOL), dimethyl ether (DME), acetone (AC) and dimethyl sulfoxide (DMSO). In order to capture the small dipole moment of TOL and also the quadrupole related with the phenyl ring, an all atom description if this molecule based on the OPLS force field was employed.[13] We note that in this molecule the multibody interactions are mainly due to the relative orientation of the rings. Atomistic DME was based on the TraPPE force field,[14,15] that is optimized to reproduce the vapor liquid equilibria. The atomistic model of AC was based on an optimized potential parametrized to reproduce liquid vapor coexistence curve, critical parameters, and vapor pressures.[16] The atomistic model of DMSO is based on an improved version of the GROMOS96 forcefield optimized for reproducing density and heat of vaporization.[17]

6.3.2 Coarse-grained model

In order to study the limit of applicability of pair potentials to describe systems where multibody contributions are important all the molecules are mapped as sin-

gle interaction site (see fig.6.2), using their centers of mass as mapping points. We note that this is an extreme choice, and we point out that in order to eventually improve the quality of the model it is possible to employ a lower degree of coarse graining.[11,12] Non bonded interaction potentials have been developed for the different molecules using the thermodynamic cycle approach (see section 6.2). Two sets of potentials have been developed for every molecule, the first one is developed carrying the pulling between the two molecules in bulk liquid, while for the second one the pulling is done in vacuum. In this latter case, since the two molecules are mapped as single beads and there are no additional contributions, the interaction potential is directly the potential of mean force ($RW(r)$) between the two molecules.

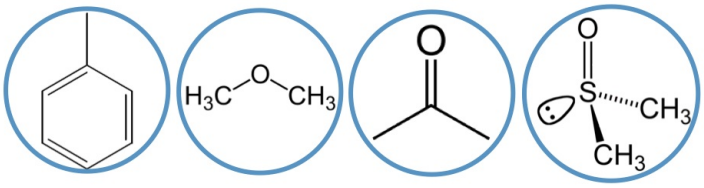


Figure 6.2: All molecules considered in this work are mapped as single interaction site in the CG representation.

6.4 Computational details

All simulations performed in this work were done using the GROMACS simulation package.[18]

6.4.1 CRW calculations

The first set of interaction potentials was developed in vacuum through a series of distance constraint simulations between the centers of mass of the molecules, ranging from 0.32 nm to 1.30 nm in steps of 0.02 nm. The constraint algorithm employed is the linear constrain solver (LINCS).[19] Lennard Jones and coulombic interactions were treated using a simple cut off, the cut off was set big enough to allow all the atoms of the two molecules to fully interact with all the others (i.e. 3.0 nm). The average constraint force is obtained from a 200 ns simulation trajectory generated employing a time step of 1 fs. The constraint dynamics simulations are performed using a stochastic dynamics integrator with an inverse friction coefficient 0.5 ps. The sampling temperature for the different molecules were: 298 K for TOL, 210 K for

DME, 270 K for AC, and 370 K for DMSO. The reversible work for every interaction is calculated integrating the average constraint force of every simulation over the constrain distance, and it can be employed directly as tabulated potential. At distance shorter than 0.32 nm the potential is extrapolated using spline interpolation.
The second sets of interaction potentials was developed employing the thermodynamic scheme presented in section 6.2. The upper part of the cycle was calculated by Boltzmann-inverting[20] the radial distribution function (RDF) between the center of mass of the molecules of NVT simulations of bulk liquid each one composed of 1000 molecules. For every system the volume was fixed in order to reproduce the average density predicted by every model at the development temperature. As before the development temperature were: 298 K for TOL, 210 K for DME, 270 K for AC, and 370 K for DMSO; the temperature were kept constant employing a velocity rescale thermostat.[21] The Lennard Jones interactions for all the systems were treated with a singel cutoff at 1.3 nm. Electrostatic interaction for TOL were treaded employing Reaction field method[22] with a reaction field cut off of 1.0 nm and outside this a considered medium dielectric constant of 2.38. For the other three molecules a PME treatment of the electrostatic[23,24] was employed, with a cutoff of 1.0 nm. In all the cases a leap frog integrator was used to update the atom positions with a time step of 1 fs. The lower part of the thermodynamic cycle was calculated employing constrain distances simulations. The constrain forces was calculated between two molecules kept at constant distance by a LINCS algorithm; as in the case of vacuum derived potentials the considered distances were raging from 0.32 nm to 1.30 nm in steps of 0.02 nm. The other simulations parameters are identical to the ones described for the calculation of $RW(r)$. Also these CG interaction potentials are used as tabulated potential in the simulations, and for distances closer than 0.32 nm the potential is extrapolated using spline interpolation.

6.4.2 Comparison between atomistic and CG models

The comparison between properties of the parents atomistic models and derived CG ones is carried considering the predicted density, the pair liquid structure and the solvation free energy calculated at the CG models development temperatures.
The density and the pair structure of the atomistic models were calculated employing NPT simulations of a 1000 molecules. The parameter of these simulations are similar to the one used for the calculations of $RW(r)$ in the previous subsection, the only difference is that this time a Berendsen barostat[25] with a coupling time of 1 ps

was employed to keep the system at a pressure of 1 atm. Also CG simulation were carried on system composed of 1000 molecules. Liquid pair structures of the CG models were calculated in NPT and in NVT conditions, in the last case the volume was fixed in order to reproduce the density predicted by the atomistic model. In all CG simulations a stochastic dynamic integrator was employed with a time step of 2 fs and an inverse friction coefficient of 0.1 ps. The cutoff of the interaction was of 1.3 nm. Since all the CG models were single bead model no coulomb interactions were present in the systems. In the CG NPT simulations the pressure was kept constant at 1 atm by a Berendsen barostat[25] with a coupling constant of 1 ps.
Thermodynamic integration[26] (TI) has been employed in order to compute the solvation free energy. In particular for every model a serie of 20 simulations with equally spaced value of the coupling parameter λ has been carried, according to wich non bonded interaction between different molecules are gradually switched on. Each simulation was 1 ns long, and the volume was fixed at the volume predicted by the atomistic model. In the simulations with atomistic models, in order to avoid singularities at small value of λ, soft core potentials have been employed;[27] this precaution is not necessary for CG models since the extrapolation of the interaction potential at small distance do not lead to a diverging value of the energy in this region. All other simulation parameters are similar to the one discussed for the calculation of the densities and the RDFs both for atomistic and CG systems.

6.5 Results and discussion

We studied the ability of CG models of a serie of molecular liquids of increasing polarity to predict properties at the state point characterizing their development (*representability*). To this end, structural and thermodynamic properties are calculated for single bead representation of: TOL, DME, AC and DMSO.

6.5.1 CG potentials

Using the CRW method we obtained CG non-bonded interaction potentials for the four different liquids. Two models have been developed for every one of them. The fist model was developed sampling the interaction between the two molecules in vacuum (vac), while the second one was developed sampling the same interaction in bulk liquid (liq). The potentials obtained for the four systems are reported in fig.6.3.

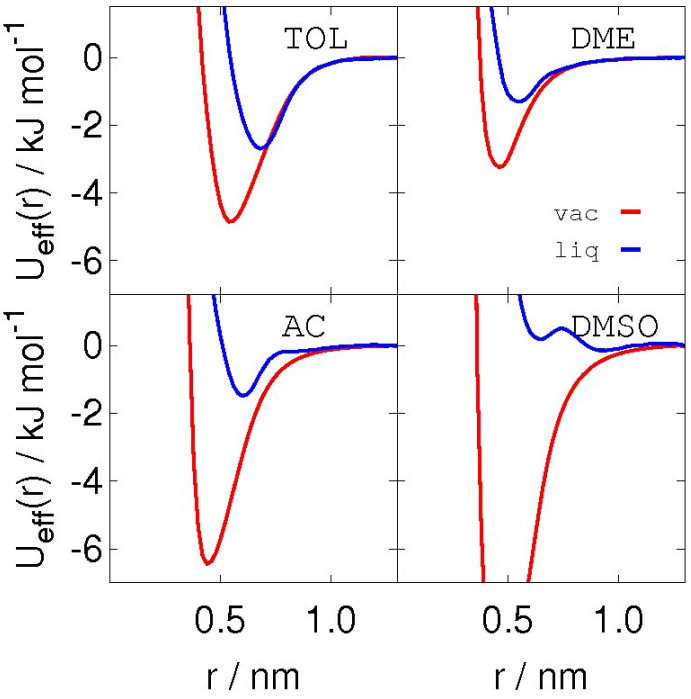


Figure 6.3: Non-bonded interaction potentials developed in this work for 1 bead model of toluene (TOL), dimethyl ether (DME), acetone (AC), and dimethyl sulfoxide (DMSO). For every model two interaction potentials have been computed, the first one is obtained carrying the sampling in vacuum (red) and the second one is obtained carrying the sampling of the interactions in bulk liquid (blue).

As it is possible to notice there is a big difference between the potentials developed in vacuum and in liquid. This is firstly imputable to the different relative configurations sampled by the molecules in the two situations. For example in the case of TOL when the sampling is carried in vacuum the two benzene ring can approach face to face, and this leads to a very strong favorable interaction; instead when the sampling is carried in bulk liquid the face to face configurations becomes a lot less favorable due to the entropic cost that this configuration has on the necessary reorganization of the surrounding molecules, and therefore its contribution to the potential is hardly sampled. This causes an evident shifting in the position and the depth of the minima of the CG interaction potential. This happens also in the case of DME, where the favorable configurations sampled at short distances by two molecules in vacuum are entropically penalized in liquid bulk. In principle this effect is present also for AC and DMSO, but in these two last cases on the top of that there are some technical problems relative to the sampling of the interactions in liquid. This difficulties are evident from the shape of the potentials, in fact as we said CRW delivers pair potentials

and therefore in principle there is no reason for obtaining interaction potentials that does not show a monotonically varying tail. This is confirmed by all the interaction potentials developed in vacuum, where the interactions are clearly pair interactions since there are no other elements in the system except the two molecules. On the contrary the interaction potentials of AC and DMSO derived in liquid show several kinks. We explain this behavior considering fig.6.4.

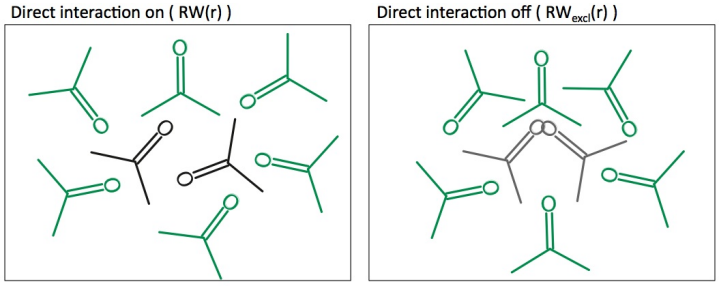


Figure 6.4: Sampling issue associated with the calculation of the potential at short distances. When the lower part of the thermodynamic cycle of fig.6.1 is calculated it is possible for the two molecules that do not interact directly to assume some unphysical configurations (right panel). These configurations can lead to strong interactions with the solvent (in this example a strong negative charge is obtained by summing the two partial charges of the oxygens). The contributions of this unphysical but extremely favorable configurations enters in $RW_{excl}(r)$, that is subtracted from $RW(r)$ in order to obtain the interaction potential. This therefore results in potentials that are less attractive than expected or in the most severe cases this can lead to the development of a completely repulsive potentials.

In the picture it is drawn a sketch of possible configurations of the molecules when the direct interactions are on ($RW(r)$) and when they are off ($RW_{excl}(r)$). In the latter case unphysical configurations can be sampled, and the interactions of those with the surrounding solvent can be favorable. These contributions enter in the reversible work calculated in the lower part of the cycle, leading to unphysical interaction potentials. This is clearly the case for DMSO, where the interaction potential obtained by liquid sampling is completely repulsive.

About the interaction potentials developed in vacuum we can say from their shape that they all represent pair interactions, but also we can notice that all these interaction potentials present deep minimas. Such potentials if employed in simulation will lead to cristal like systems. This results is somehow expected since these systems have strong directional interactions, that are “counterbalanced” by the entropy of the surrounding fluid. Since this is not present in the vacuum sampling the obtained interaction potentials can not properly describe a liquid state.

	TOL		DME		AC	
	AA	CG	UA	CG	UA	CG
ρ / $kg\ m^{-3}$	871.6	641.1	785.7	406.5	833.6	156.9
ΔG / $kJ\ mol^{-1}$	-18.3	-6.3	-11.2	-2.9	-18.0	+2.0

Table 6.1: Comaprison between the thermodynamic properties (density and solvation free energy) predicted by the CG models developed in liquid and their parent atomistic model.

The potentials developed in liquid, excluding DMSO, have more the characteristics of potentials of single beads fluid. Still it has to be kept in mind that we are developing spherically symmetric interaction potentials for systems with directional interactions and that some artifacts can be present due to sampling issue.
In the next subsection we compare the structure and some thermodynamic properties of the CG models developed in liquid with the ones predicted by the parents atomistic models, trying to understand to which extent pair potentials can be used to simulate system where multibody interactions are important.

6.5.2 Representability of the CG models

In tab.6.1 are reported the investigated thermodynamic properties predicted by the CG and the parent atomistic models. The densities were calculated at 1 atm at the temperature used during the development of the CG models, the solvation free energy were also calculated at the temperature used in the development of the CG model, but the densities of the CG calculation were set to be the same as the ones of the atomistic models.

As it is possible to see the density are underestimated, and the representability of the model decreases as the severity of the assumption of pair interaction increases (i.e. the polarity of the molecules increases). The solvation free energy is an extremely sensitive property to the quality of the interaction potentials.[28,29] No one of the models is really able to predict it, this is an indication of the poorness of the chosen spherical symmetry of the interactions. The fact that the CG model of AC predicts a solvation free energy with the wrong sign points out that not only the assumption of pair additivity of potential is a poor assumption for molecule with a strong dipole moment, but also it confirms that already for this system there are severe sampling issue in the calculation of $RW_{excl}(r)$.
In fig.6.5 are reported the radial distribution functions of the three systems calculated at the atomistic level (black lines), an at the CG level. Of this a first series of

calculations is carried NPT at the pressure of 1 atm (geen), and a second series is carried NVT at the density predicted by the atomistic models (blue). As expected form the density data the CG NPT simulations predict the positions of the peaks shifted to bigger distances. Also they predict less structured RDFs. NVT simulations of TOL and DME show a good agreement between the structure predicted by the CG models and the one of the parent atomistic ones. The position of the first peaks their shapes ae predicted correctly by both models, the postion of the second peaks and their shapes are also predicted reasonably correctly. The CG NVT simulation of AC show that the model is able to recover the postion of the first peak, but not completely its correct shape. This is probably related with sampling issue that affects the quality of the interaction potentials.

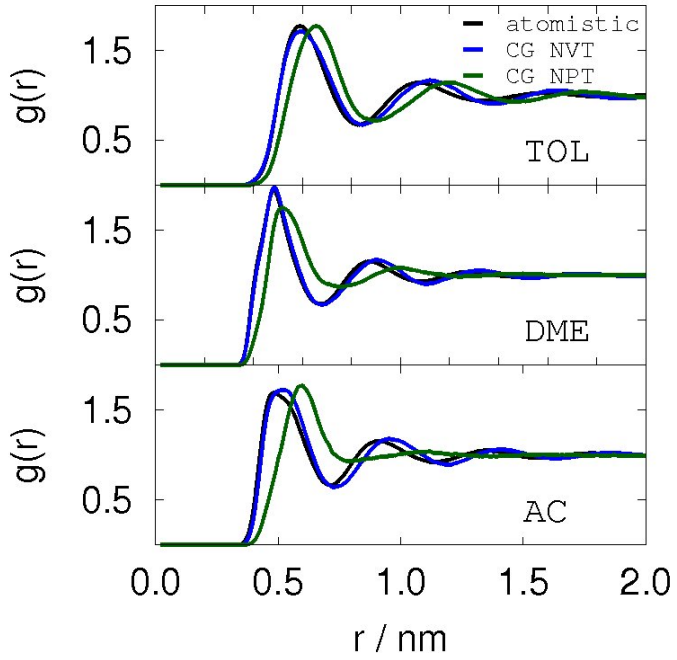


Figure 6.5: Comparison of the RDFs between the reference atomistic simulations and CG ones. The CG simulations have been carried both in NPT and in NVT condition at the density predicted by the atomistic models. CG models underestimate the density of the systems, therefore in NPT condition their peaks appear shifted. In NVT condition the postions and the shapes of the peaks are recovered for TOL and DME. The postions of the peaks are also partially recovered for AC, but not their shapes.

6.6 Conclusions

We have shown that in systems dominated by multibody interactions the assumption to use pair potentials to describe the interaction between molecules leads to models with a poor representability. On the top of that we have shown that for highly polar molecular liquids the thermodynamic cycle employed to calculate CRW potentials can lead to sampling problem.

In order to recover the first limitation it is possible to employ a lower level of CG of the system in order to retain information about directional interaction in the mapping scheme. An example of this is shown in the work of Brini et al.[12], where a molecule of toluene is coarse grained using a 3 beads mapping. This enables the possibility of maintaining topological informations into the model, that then is able to express a good representability.

A possible way to solve the sampling issue is to employ different methods to calculate the interaction free energy between the groups of atoms (e.g. TI or perhaps thermodynamic perturbation). In alternative it is in principle possible to constrain the relative orientation of the molecules in order to avoid the sampling of unphysical configurations during the calculation of $RW_{excl}(r)$.

6.7 Bibliography

[1] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. *J. Phys. Chem. B*, 111:7812, 2007.

[2] J. D. McCoy and J. G. Curro. *Macromolecules*, 31:9362, 1998.

[3] B. M. Mognetti, P. Virnau, L. Yelash, W. Paul, K. Binder, M. Müller, and L. G. MacDowell. *J. Chem. Phys.*, 130:044101, 2009.

[4] D. Reith, M. Pütz, and F. Müller-Plathe. *J. Comput. Chem.*, 24:1624–1636, 2003.

[5] A. Lyubartsev and A. Laaksonen. *Phys. Rev. E*, 52:3730, 1995.

[6] F. Ercolessi and J. Adams. *Europhys. Lett.*, 26:583, 1994.

[7] M. S. Shell. *J. Chem. Phys.*, 129:144108, 2008.

[8] S. Izvekov and G. A. Voth. *J. Phys. Chem. B*, 109:2469, 2005.

[9] J. W. Mullinax and W. G. Noid. *J. Phys. Chem. C*, 114:5661, 2010.

[10] B. Hess, C. Holm, and N. F. A. van der Vegt. *J. Chem. Phys.*, 124:164509, 2006.

[11] Y. Wang, W. G. Noid, P. Liu, and G. A. Voth. *Phys. Chem. Chem. Phys.*, 11:2002, 2009.

[12] E. Brini, V. Marcon, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 13:10468, 2011.

[13] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. *J. Am. Chem. Soc.*, 118:11225, 1996.

[14] M. G. Martin and J. I. Siepmann. *J. Phys. Chem. B*, 102:2569, 1998.

[15] J. M. Stubbs, J. J. Potoff, and J. I. Siepmann. *J. Phys. Chem. B*, 108:17596, 2004.

[16] G. Kamath, G. Georgiev, and J. J. Potoff. *J. Phys. Chem. B*, 109:19463–19473, 2005.

[17] D. P. Geerke, C. Oostenbrink, N. F. A. van der Vegt, and W. F. van Gunsteren. *J. Phys. Chem. B*, 108:1436, 2004.

[18] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Theory Comput.*, 4:435, 2008.

[19] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. *J. Comput. Chem.*, 18:1463, 1997.

[20] W. Tschoep, K. Kremer, J. Batoulis, T. Burger, and O. Hahn. *Acta Polym.*, 49:61, 1998.

[21] G. Bussi, D. Donadio, and M. Parrinello. *J. Chem. Phys.*, 126:014101, 2007.

[22] J. A. Barker and R. O. Watts. *Mol. Phys.*, 26:789, 1973.

[23] T. Darden, D. York, and L. Pedersen. *J. Chem. Phys.*, 98:10089, 1993.

[24] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. *J. Chem. Phys.*, 103:8577, 1995.

[25] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. *J. Chem. Phys.*, 81:3684, 1984.

[26] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Academic Press, 2nd edition, 2001.

[27] B. Hess and N. F. A. van der Vegt. *J. Phys. Chem. B*, 110:17616, 2006.

[28] J. W. Shen, C. Li, N. F. A. van der Vegt, and C. Peter. *J. Chem. Theory Comput.*, 7:1916, 2011.

[29] E. Brini, C. R. Herbers, G. Deichmann, and N. F. A. van der Vegt. *Phys. Chem. Chem. Phys.*, 14:11896, 2012.

7 Outlook

This thesis presented a novel method to systematically coarse grain molecules: the conditional reversible work (CRW) method. This defines the interaction potentials between the beads as the free energy calculated between the groups of atoms that are represented by the beads under the condition that they are embedded in their respective molecules. We have shown that three are the key characteristics of the potentials developed using this method: they are pair potentials, they have a clear link with a physical quantity and they represent conditional interactions. The fact that they represent pair interaction free of state point dependent multibody contributions on one side makes the model that uses this potential more transferable, but on the other side this requires more carefulness when coarse grained (CG) models are build for systems where multibody contributions are important. The clear link to a physical quantity enables the opportunity to understand and eventually repair failure of the models. The fact that the potentials are representative of conditional interactions enables the possibility to employ these potentials in a building block kind of approach to build CG models for big molecules. This thesis highlights all these aspects providing a solid first step in the use of the CRW method for the coarse graining of soft matter systems.

Further routes can be foreseen in the application of CRW methods.
The first one it is simply the use of CRW as CG method to build simplified models for van der Waals dominated systems. Of particular interest can be the creation of CG models for tecnologically relevant polymers like poly ethylene, poly propylene, poly butadiene, poly tetrafluoroethylene.
A more challenging route to follow is the CG of systems characterized by hydrophobic interactions. These are crucial to determine the behavior of macromolecules in water and therefore the development of CG techniques able to properly represent these systems is extremely appealing, and they also have pairwise characteristics. CRW is able in principle to deliver pair potentials that can capture the interaction free energy of between the different groups in water, and even more important this free energy will be a *conditional* one. This means that into the interaction potentials

there will be information about the solvation of the groups when they are embedded in their molecules. This is of extreme importance since it is easy to imagine that the organization of the solvation sphere around a group of atoms is different if the group stands alone or if it is embedded in a bigger molecule. This is most likely to be a strong improvement over the so called fragment based CG, where CG models are built using as building blocks fragment of molecules that assumes that the difference between the solvation of an isolated molecule block and the same block embedded in a molecule is negligible. The biggest challenge in this approach consists in identifying a cheap method to calculate the interaction free energy, since both the methods we proposed are likely to fail due to sampling issue or to too big perturbations and alternative methods to calculate free energy (e.g. thermodynamic integration) are too expensive.

The peculiar characteristics of the CRW developed potentials can be also used to gain a better understanding of the effect that the CG process has on the systems. For example, since CRW delivers interaction potentials with a clear physical meaning and free of indirect contributions, it is in principle possible to employ it to systematically study the effects of different mapping schemes on the thermodynamic and dynamic properties of a given system. An other possibility enabled by the pairwise nature of the CRW interaction potentials is the definition of a pairwise friction between the beads. This property can be calculated simultaneously to the computation of the interaction free energy, and in principle can be used as input in specific thermostat (e.g. dissipative particle dynamics). This can lead to thermodynamic models able to reproduce not only the thermodynamic of the systems but also their dynamic, allowing a consistent study of non equilibrium phenomena with CG models.

Curriculum Vitae

Name: Emiliano Brini

Place of Birth: Bologna (Bo), Italy

Date of Birth: 05.12.1983

Nationality: Italian

Education:

1997 – 2002: High school diploma: Istituto tecnico industriale Aldini-Valeriani, (Bologna, Italy)

2002 – 2006: Bachelor: Universitá degli studi di Bologna (Bologna, Italy)

2006 – 2008: Master: Universitá degli studi di Bologna (Bologna, Italy)

2009 – 2012: Center of Smart Interfaces, Darmstadt University of Technology

_____

(Emiliano Brini)

Emiliano Brini
Dolivostr. 21
64293 Darmstadt

Tel.: 015152593710
E-mail: brini@csi.tu-darmstadt.de

Darmstadt, den 24. August 2012

Erklärung

Ich erkläre hiermit, noch keinen Promotionsversuch unternommen zu haben.

_____

(Emiliano Brini)

Emiliano Brini
Dolivostr. 21
64293 Darmstadt

Tel.: 015152593710
E-mail: brini@csi.tu-darmstadt.de

Darmstadt, den 24. August 2012

Eidesstattliche Erklärung

Ich erkläre mich hiermit an Eides Statt, dass ich meine Dissertation selbständig und nur mit den angegebenen Hilfen angefertigt habe.

_____

(Emiliano Brini)